

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 10:14:07 ; Search time 1391.78 Seconds
(without alignments)
11437.798 Million cell updates/sec

Title: US-10-660-763-1

Perfect score: 2674
Sequence: 1 tccgggggtccgcacccgggccc.....aaaaaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5544816 seqs, 297611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2674	100.0	2674	13	US-10-003-295-1 Sequence 1, Appl1
2	2674	100.0	2674	17	US-10-660-763-1 Sequence 1, Appl1
3	2326.4	87.0	2889	16	US-10-240-965-256 Sequence 256, App
4	362.4	13.6	15297	13	US-10-003-295-3 Sequence 3, Appl1
5	362.4	13.6	15297	17	US-10-660-763-3 Sequence 3, Appl1
6	326.8	12.2	449	10	US-09-918-995-1503 Sequence 1503, Ap
7	326.8	11.1	361	9	US-09-948-802-5 Sequence 5, Appl1
8	296.2	11.1	361	15	US-10-121-925-5 Sequence 19, Appl1
9	271.6	10.2	1779	17	US-10-280-576-19 Sequence 19, Appl1
10	201.2	7.5	3875	15	US-10-101-510-525 Sequence 525, App
11	198	7.4	2955	18	US-10-384-339C-1 Sequence 1, Appl1

12	198	7.4	3370	9	US-09-967-768A-144 Sequence 144, App
13	198	7.4	3370	16	US-10-354-358-101 Sequence 101, App
14	198	7.4	3370	18	US-10-210-120-19 Sequence 19, Appl1
15	198	7.4	3370	18	US-10-776-827-82 Sequence 82, Appl1
16	198	7.4	3370	18	US-10-473-974-219 Sequence 219, App
17	185.4	6.9	3042	18	US-10-384-339C-2 Sequence 2, Appl1
18	185.4	6.9	3921	10	US-09-921-406C-29 Sequence 29, Appl1
19	185.4	6.9	3921	15	US-10-007-926A-222 Sequence 222, App
20	185.4	6.9	3921	16	US-10-269-909-19 Sequence 19, Appl1
21	185.4	6.9	3921	17	US-10-366-288-81 Sequence 1, Appl1
22	185.4	6.9	3921	17	US-10-172-118-1036 Sequence 1036, Ap
23	185.4	6.9	3921	17	US-10-342-887-1036 Sequence 1036, Ap
24	185.4	6.9	3921	17	US-10-287-226-319 Sequence 319, App
25	185.4	6.9	3921	18	US-10-648-593-1 Sequence 1, Appl1
26	185.4	6.9	3921	18	US-10-473-974-223 Sequence 223, App
27	185.4	6.9	3921	19	US-10-616-403-5 Sequence 5, Appl1
28	185.4	6.9	3935	19	US-10-897-711-1 Sequence 1, Appl1
29	176.2	6.6	1518	17	US-10-280-576-24 Sequence 24, Appl1
30	174.6	6.5	816	18	US-10-377-268-3 Sequence 3, Appl1
31	174.6	6.5	1050	18	US-10-377-268-4 Sequence 4, Appl1
32	174.6	6.5	3030	16	US-10-325-430-8 Sequence 8, Appl1
33	174.6	6.5	3416	8	US-08-987-689A-1 Sequence 1, Appl1
34	174.6	6.5	3416	15	US-10-292-524-1 Sequence 1, Appl1
35	174.6	6.5	3416	17	US-10-464-805-2 Sequence 2, Appl1
36	174.6	6.5	3416	17	US-10-305-720-1483 Sequence 1483, Ap
37	174.6	6.5	3544	14	US-10-161-803-51 Sequence 51, Appl1
38	174.6	6.5	4089	17	US-10-291-808-13 Sequence 13, Appl1
39	174.6	6.5	4089	18	US-10-620-052A-7 Sequence 7, Appl1
40	174.6	6.5	4151	14	US-10-161-803-52 Sequence 52, Appl1
41	174.6	6.5	4151	16	US-10-325-430-7 Sequence 7, Appl1
42	170.4	6.4	1611	18	US-10-476-962-3 Sequence 3, Appl1
43	170.4	6.4	1611	18	US-10-887-588-1 Sequence 1, Appl1
44	170.4	6.4	1611	18	US-10-887-588-3 Sequence 3, Appl1
45	170.4	6.4	1611	19	US-10-497-641-1 Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-10-003-295-1
; Sequence 1, Application US/10003295
; Publication No. US20020168741A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weidun et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL0001183DIV
; CURRENT APPLICATION NUMBER: US/10/003,295
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2674
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-003-295-1

Query Match	100.0%	Score 2674	DB 13	Length 2674
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2674	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	TCGGGGTCCGACCGGGCTGAGTGCCTCCAGAGCGAGCTGCCGCTGC	60	
DB	1	TCGGGGTCCGACCGGGCTGAGTGCCTCCAGAGCGAGCTGCCGCTGC	60	
QY	61	GGAACGACCTATGGCTTCTTCTGAGCTGTGACGCCGCCAGGGGCTCTCTGC	120	
DB	61	GGAACGACCTATGGCTTCTTCTGAGCTGTGACGCCGCCAGGGGCTCTCTGC	120	
QY	121	AGCAATGAGAGGCGGAGCTTCGTCTGAGGAGGAGTGAAGTGGATGCCAGC	180	

Db 121 AGCAAAATGAGAGGAGCCGAGCTTCCTCTACTGAGAGGAGCATGAGAAATGAGATGAGCCCAAC 180
QY 181 GGGTCAAGATGACAGAGAGTATGACGAGACTGCTTACCAATGATCCCTGAGAGAGATG 240
Db 181 GGGTCAAGATGACAGAGAGTATGACGAGACTGCTTACCAATGATCCCTGAGAGAGATG 240
QY 241 GGGGCGAGAGCCGGGCGCATCAGCCCTGACAGCCCATCATGTCAGTCCCTGGGCTGAGATCA 300
Db 241 GGGGCGAGAGCCGGGCGCATCAGCCCTGACAGCCCATCATGTCAGTCCCTGGGCTGAGATCA 300
QY 301 CCAGCCAACTGAGGGGCTGAGGCGCTGCTGCGGACGACCGAGAGATCTGAATCTCAG 360
Db 301 CCAGCCAACTGAGGGGCTGAGGCGCTGCTGCGGACGACCGAGAGATCTGAATCTCAG 360
QY 361 GGGCCCTGAGCAAGCTGAGGCTGCTCATCCGGGAGAGGGGAGGAGCTTCGAGAACCTTACA 420
Db 361 GGGCCCTGAGCAAGCTGAGGCTGCTCATCCGGGAGAGGGGAGGAGCTTCGAGAACCTTACA 420
QY 421 GCGAGAGTGGACGAGCTGACAGAGAGCTCACAAGACCCACAGCCAGACATTTGAGA 480
Db 421 GCGAGAGTGGACGAGCTGACAGAGAGCTCACAAGACCCACAGCCAGACATTTGAGA 480
QY 481 AGCTGAAGACCAAGTACCGAGCTCTGCGACCGGACAGTGCCTCAAGCCAGCAAGTACC 540
Db 481 AGCTGAAGACCAAGTACCGAGCTCTGCGACCGGACAGTGCCTCAAGCCAGCAAGTACC 540
QY 541 AGGAGGCGAGCAAGACAGAGACCGTGAACAAGGCGCAAGGACAGTATGATGCGAGCTGT 600
Db 541 AGGAGGCGAGCAAGACAGAGACCGTGAACAAGGCGCAAGGACAGTATGATGCGAGCTGT 600
QY 601 GGAAGCTCTTTGCTCACAACAACCGCTATGCTGAGGCGTGGGCGTGGGCTGAGCACTTAC 660
Db 601 GGAAGCTCTTTGCTCACAACAACCGCTATGCTGAGGCGTGGGCGTGGGCTGAGCACTTAC 660
QY 661 ACCAGACCAACAACAAGCTCTGCTGCGCGGCTGCTGCGGTAATGACAGAACTTGAAG 720
Db 661 ACCAGACCAACAACAAGCTCTGCTGCGCGGCTGCTGCGGTAATGACAGAACTTGAAG 720
QY 721 AGGAGATGAGTTCATCTGGAAGAGATGCTGAGGAATACCTGAGGAATTAAGAGCTGG 780
Db 721 AGGAGATGAGTTCATCTGGAAGAGATGCTGAGGAATACCTGAGGAATTAAGAGCTGG 780
QY 781 TGCAGATGAGTGTGAGGCTTACCGGAGATGAGCTGAGCTGCTGCGGCTGAGCAAG 840
Db 781 TGCAGATGAGTGTGAGGCTTACCGGAGATGAGCTGAGCTGCTGCGGCTGAGCAAG 840
QY 841 CTGAGGCTGAGTACCAAGGCTTCTGCGACAGTATGAGTTCGCACTGACGCTCCACCT 900
Db 841 CTGAGGCTGAGTACCAAGGCTTCTGCGACAGTATGAGTTCGCACTGACGCTCCACCT 900
QY 901 GTGTCAAGTTCGATGAGTCACTGCTTGAAGAGGAGTGAACCGCTGAGAGCTGGGAGACTCC 960
Db 901 GTGTCAAGTTCGATGAGTCACTGCTTGAAGAGGAGTGAACCGCTGAGAGCTGGGAGACTCC 960
QY 961 AGCTGAACGAGCTGATGTGAGAGAGCTGACAGACAGCTGACCTTGAAGTGAAGTGAAG 1020
Db 961 AGCTGAACGAGCTGATGTGAGAGAGCTGACAGACAGCTGACCTTGAAGTGAAGTGAAG 1020
QY 1021 TGGCTGTGCTCAGAGAGATGTGTTCAAGCGGACAGAGATGTTATGAGAGCTGCAACAG 1080
Db 1021 TGGCTGTGCTCAGAGAGATGTGTTCAAGCGGACAGAGATGTTATGAGAGCTGCAACAG 1080
QY 1081 AGCTCGGAAATGAAGAGAGAACCAACCCCGGAGAGCGGAGTGAAGCTGTGGGAGAG 1140
Db 1081 AGCTCGGAAATGAAGAGAGAACCAACCCCGGAGAGCGGAGTGAAGCTGTGGGAGAG 1140
QY 1141 GCGAAGTGTGCAAGAGACATGCAAGGAGCTGCAAGTGAAGCTGTGCAAGCAAGCAAG 1200
Db 1141 GCGAAGTGTGCAAGAGACATGCAAGGAGCTGCAAGTGAAGCTGTGCAAGCAAGCAAG 1200
QY 1201 TGCAGGCGCAGAGAGAGTGTGCTGCAAGCAAGCTGAGAGCACTGGGCGCGGAGAGCC 1260
Db 1201 TGCAGGCGCAGAGAGAGTGTGCTGCAAGCAAGCTGAGAGCACTGGGCGCGGAGAGCC 1260

QY 1261 CGCTGTGCTGCTCTGACAGATGACCGGCACTCCAGTGTCTCTGAGAGAGAGAGAG 1320
Db 1261 CGCTGTGCTGCTCTGACAGATGACCGGCACTCCAGTGTCTCTGAGAGAGAGAGAG 1320
QY 1321 AGGGGAGAGAGACACCCAGCTGAGATCTTTAAGACCAATCTCAGAAATCTTCGCC 1380
Db 1321 AGGGGAGAGAGACACCCAGCTGAGATCTTTAAGACCAATCTCAGAAATCTTCGCC 1380
QY 1381 CGAAGTCTCGAAGCTGTACCGAGCTGGAAGGGGAGGCTTCTGATGCAATCTCTTGTCTCA 1440
Db 1381 CGAAGTCTCGAAGCTGTACCGAGCTGGAAGGGGAGGCTTCTGATGCAATCTCTTGTCTCA 1440
QY 1441 TGCACCACTTACTGAGACACCCAGAGCCCTTACCAAGAAAGTGTGTTCTGCTGACA 1500
Db 1441 TGCACCACTTACTGAGACACCCAGAGCCCTTACCAAGAAAGTGTGTTCTGCTGACA 1500
QY 1501 GGGCTGTGCGCAAGAGCAAGTGGTGTGAACCATGAGGACCTGTGTGGTGTGACAGA 1560
Db 1501 GGGCTGTGCGCAAGAGCAAGTGGTGTGAACCATGAGGACCTGTGTGGTGTGACAGA 1560
QY 1561 TTGAGACGGGAGAACTTTGGCGAAGTGTTCAGGGAGAGCGCTTGGAGAGCGCAACACCTGG 1620
Db 1561 TTGAGACGGGAGAACTTTGGCGAAGTGTTCAGGGAGAGCGCTTGGAGAGCGCAACACCTGG 1620
QY 1621 TGGCGGTGAAGTCTTGTGAGAGAGCGCTCCACTGAACTTCAAGGCGCAAGTTTCTACAG 1680
Db 1621 TGGCGGTGAAGTCTTGTGAGAGAGCGCTCCACTGAACTTCAAGGCGCAAGTTTCTACAG 1680
QY 1681 AAGCAGAGATCTTGAAGAGATACAGCAACCCCAACATGCTGTGCTCTCATTTGTGTGCA 1740
Db 1681 AAGCAGAGATCTTGAAGAGATACAGCAACCCCAACATGCTGTGCTCTCATTTGTGTGCA 1740
QY 1741 CCCAAGAGAGGCGCAATCTACATGCTCATGAGACTTGTGAGAGGGGCGGCACTTCTGACCT 1800
Db 1741 CCCAAGAGAGGCGCAATCTACATGCTCATGAGACTTGTGAGAGGGGCGGCACTTCTGACCT 1800
QY 1801 TCCTCCGACAGAGAGGAGGCGCGCTGCGGAGTGAAGCTGTGCTGAGATGATGAGGAGATG 1860
Db 1801 TCCTCCGACAGAGAGGAGGCGCGCTGCGGAGTGAAGCTGTGCTGAGATGATGAGGAGATG 1860
QY 1861 CAGCTGTGAGTGAAGTACTGAGAGAGCAAGTGTGCTGACCAACCGGAGACTGCGCTC 1920
Db 1861 CAGCTGTGAGTGAAGTACTGAGAGAGCAAGTGTGCTGACCAACCGGAGACTGCGCTC 1920
QY 1921 GGAATGCTGCTGAGTGAAGAGAGATGCTGTAAGATGAGTGAATGAGTGTGAGATG 1980
Db 1921 GGAATGCTGCTGAGTGAAGAGAGATGCTGTAAGATGAGTGAATGAGTGTGAGATG 1980
QY 1981 AGGAAAGCCGATGAGGCTTATGAGAGCTTCAAGGAGGCTCAGACAAAGTCCCGTGAAGTGA 2040
Db 1981 AGGAAAGCCGATGAGGCTTATGAGAGCTTCAAGGAGGCTCAGACAAAGTCCCGTGAAGTGA 2040
QY 2041 CCGCACTTGAAGGCGCTTAACTACGCGCGCTACTCTCCGAAAGCGAGCTGTGAGAGCTTTG 2100
Db 2041 CCGCACTTGAAGGCGCTTAACTACGCGCGCTACTCTCCGAAAGCGAGCTGTGAGAGCTTTG 2100
QY 2101 GGAATGCTGCTGAGAGAGCTTCAAGCTGAGGAGGCTCCCTATGCCAATCTCAGAAATC 2160
Db 2101 GGAATGCTGCTGAGAGAGCTTCAAGCTGAGGAGGCTCCCTATGCCAATCTCAGAAATC 2160
QY 2161 AGCAGACAGGAGAGTGTGAGAGAGGAGGAGGCGCTGTGCTGCTGCGCAAGAGCTGTGCTG 2220
Db 2161 AGCAGACAGGAGAGTGTGAGAGAGGAGGAGGCGCTGTGCTGCTGCGCAAGAGCTGTGCTG 2220
QY 2221 ATGCGGTGTTCAAGCTCATGAGAGAGCTGTGAGGCTGTAGAGCTGTGAGAGCTGTGAG 2280
Db 2221 ATGCGGTGTTCAAGCTCATGAGAGAGCTGTGAGGCTGTAGAGCTGTGAGAGCTGTGAG 2280
QY 2281 TGCAGCACTTACCAAGAGAGCTGCAAGAGATCCGAAAGGCGGCAATCGGTGAGGCTGAGAGC 2340
Db 2281 TGCAGCACTTACCAAGAGAGCTGCAAGAGATCCGAAAGGCGGCAATCGGTGAGGCTGAGAGC 2340

[illegible]

```

RESULT 2
US-10-660-763-1
/ Sequence 1, Application US/10660763
/ Publication No. US20040063130A1
/ GENERAL INFORMATION:
/ APPLICANT: GAN, Weidun et al.
/ TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
/ TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
/ TITLE OF INVENTION: THEREOF
/ FILE REFERENCE: C0001183DIV11
/ CURRENT APPLICATION NUMBER: US/10/660,763
/ CURRENT FILING DATE: 2003-09-12
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 2674
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-660-763-1

```

Query Match	100.0%;	Score 2674;	DB 17;	Length 2674;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2674; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	TCGGGGGCTCCGACACCGGGCCCTGAGTCCGAGGCGCTCCAGAGACAGTGCCTCGC	60
Db	1	TCGGGGGCTCCGACACCGGGCCCTGAGTCCGAGGCGCTCCAGAGACAGTGCCTCGC	60
QY	61	GGAA CAGC ACTATGGGCTTCTCTTCTGAGCTGTGAGCCCCAGGGCACGGGGTCTCGC	120
Db	61	GGAA CAGC ACTATGGGCTTCTCTTCTGAGCTGTGAGCCCCAGGGCACGGGGTCTCGC	120
QY	121	AGCAAAATGCAAGAGGCGGAGCTTCGTCTACTGAGAGGCGATGAGAAATGATGGCCCCAGC	180
Db	121	AGCAAAATGCAAGAGGCGGAGCTTCGTCTACTGAGAGGCGATGAGAAATGATGGCCCCAGC	180
QY	181	GGGTCGAAGAGTGCACAGGAGATATGCAAGAGCTGCTTCACCA C ATGCTCCCTGACGAGCACTG	240
Db	181	GGGTCGAAGAGTGCACAGGAGATATGCAAGAGCTGCTTCACCA C ATGCTCCCTGACGAGCACTG	240
QY	241	GGGGCCAGAGCGGGCCATTCAGCCCTGACAGCCCATATGTCATGCTCCGGGCTGAGATCA	300
Db	241	GGGGCCAGAGCGGGCCATTCAGCCCTGACAGCCCATATGTCATGCTCCGGGCTGAGATCA	300
QY	301	CCAGCCAAATGAGAGGCGCTGAGCGGCTTGCTGCGGACAGCAGAGGATCTGAACTCAG	360
Db	301	CCAGCCAAATGAGAGGCGCTGAGCGGCTTGCTGCGGACAGCAGAGGATCTGAACTCAG	360

Qy	361	GGCCCCGAGACACTGACCTGCAATCCGGGAAACGGACAGACCTTCGAAAGACTCA	420
Dp	361	GGCCCCGAGACACTGACCTGCAATCCGGGAAACGGACAGACCTTCGAAAGACTCA	420
Qy	421	GCGAGCAGTGGCAGCAGCTGCAGAGAGCTCAACCAAGCCACAGCCAGGACATTGAGA	480
Dp	421	GCGAGCAGTGGCAGCAGCTGCAGAGAGAGCTCAACCAAGCCACAGCCAGGACATTGAGA	480
Qy	481	AGCTGAAGACCAATACCGAGCTCTGSCACGGGACAGTGGCCCAAGCCAAAGCGAATACC	540
Dp	481	AGCTGAAGACCAATACCGAGCTCTGSCACGGGACAGTGGCCCAAGCCAAAGCGAATACC	540
Qy	541	AGAGAGCCACGAAAGACAGGACCTGTGACAAAGGCCAAAGACAAATATGTGGGACCTGT	600
Dp	541	AGAGAGCCACGAAAGACAGGACCTGTGACAAAGGCCAAAGACAAATATGTGGGACCTGT	600
Qy	601	GGAAGCTCTTTGCTGCACCAACACCGCTATGTCGTGGGCGTGGGGCTGGCAGGTCACAC	660
Dp	601	GGAAGCTCTTTGCTGCACCAACACCGCTATGTCGTGGGCGTGGGGCTGGCAGGTCACAC	660
Qy	661	ACCAGCACACACACAGCTCTGTGTGCCGGGCTGTGCGGTCACTGACGAGACCTGCACG	720
Dp	661	ACCAGCACACACACAGCTCTGTGTGCCGGGCTGTGCGGTCACTGACGAGACCTGCACG	720
Qy	721	AGGAGATGGCTTGATCCTGAAAGAAATCCTGCAGGAATATCTGGAAGTTATGACGCTGG	780
Dp	721	AGGAGATGGCTTGATCCTGAAAGAAATCCTGCAGGAATATCTGGAAGTTATGACGCTGG	780
Qy	781	TGCGAGATGAGTGGTGGCAATTCACGGGGAGATGGCTGACGCTGGCCCGCATTCAGC	840
Dp	781	TGCGAGATGAGTGGTGGCAATTCACGGGGAGATGGCTGACGCTGGCCCGCATTCAGC	840
Qy	841	CTGAGGCTGAGTACCAAGGCTTCTGTGCACAGTATGGGTCCGCACTGACCTCCACCCT	900
Dp	841	CTGAGGCTGAGTACCAAGGCTTCTGTGCACAGTATGGGTCCGCACTGACCTCCACCCT	900
Qy	901	GTCGTACGTTGATGAGTCACTGCTTTGAGAGGGGTAAACCGCTGAGCTTGGGGAGCTCC	960
Dp	901	GTCGTACGTTGATGAGTCACTGCTTTGAGAGGGGTAAACCGCTGAGCTTGGGGAGCTCC	960
Qy	961	AGCTGAACGAGCTGACCTGTGGAGGCGGTGACGACACGCTGACCTCGATGACATGAGC	1020
Dp	961	AGCTGAACGAGCTGACCTGTGGAGGCGGTGACGACACGCTGACCTCGATGACATGAGC	1020
Qy	1021	TGCGCTGTGGCCACCGAATGGTGTTCAGGGCGGACGAGATGGTTATCCGACGTGCACACG	1080
Dp	1021	TGCGCTGTGGCCACCGAATGGTGTTCAGGGCGGACGAGATGGTTATCCGACGTGCACACG	1080
Qy	1081	AGCTCCGGAATGAAGAGAGAAACCCACCCCGGGAGCGGGTGCACGTCTGGGGCAAGA	1140
Dp	1081	AGCTCCGGAATGAAGAGAGAAACCCACCCCGGGAGCGGGTGCACGTCTGGGGCAAGA	1140
Qy	1141	GGCAAGTGTCTCAAGAGCACTGAGAGGGGCTGCAGGTATAGGGCTGTGCACCGACCAAGC	1200
Dp	1141	GGCAAGTGTCTCAAGAGCACTGAGAGGGGCTGCAGGTATAGGGCTGTGCACCGACCAAGC	1200
Qy	1201	TGCAAGGCTCCAGAGGAGTGTGCTGCAGACCAAGCTGAGAGCACTTGGGCCCGGCAAGCCC	1260
Dp	1201	TGCAAGGCTCCAGAGGAGTGTGCTGCAGACCAAGCTGAGAGCACTTGGGCCCGGCAAGCCC	1260
Qy	1261	CGCCTGTGTGCTCTCTGACAGATGACCGCACTTCACGTGTCTCTGGAGCAGAGCGAG	1320
Dp	1261	CGCCTGTGTGCTCTCTGACAGATGACCGCACTTCACGTGTCTCTGGAGCAGAGCGAG	1320
Qy	1321	AGGGGGGAAAGACCCACGCTGAGATCTTTAAGACCAATTCAGGAATCTTCCGCC	1380
Dp	1321	AGGGGGGAAAGACCCACGCTGAGATCTTTAAGACCAATTCAGGAATCTTCCGCC	1380
Qy	1381	CCAAGTCTGGAACCTGTATCCGATGGAAGGGGAAAGGCTTCTTCAAGCATTCCTTTGCTCA	1440
Dp	1381	CCAAGTCTGGAACCTGTATCCGATGGAAGGGGAAAGGCTTCTTCAAGCATTCCTTTGCTCA	1440
Qy	1441	TCGACCAACTTACTGAGCACCCAGCAGCCCTTCACCAAGAAAGTGTGTCTCTGCACA	1500

```

Db      1441 TGACACCACTACTGACACACCCAGACCCCTCAACAAGAAAGATGATGTTGCTCTGACAA 1500
Qy      1501 GGGCTGTGCCCAAGAGCAAGTGGGTCTGAACCATAGAGCACTGTGTGTTGGGTGACACA 1560
Db      1501 GGGCTGTGCCCAAGAGCAAGTGGGTCTGAACCATAGAGCACTGTGTGTTGGGTGACACA 1560
Qy      1561 TTGGACGGGGGAACTTTGGCGAAGTGTTCAGCGGACGCTGCGAGCCGAACAACCCCTG 1620
Db      1561 TTGGACGGGGGAACTTTGGCGAAGTGTTCAGCGGACGCTGCGAGCCGAACAACCCCTG 1620
Qy      1621 TGGCGGTGAAGTCTTGTGAGAGACGCTCCACCTGACCTCAAGGCAAGTTTCTACAG 1680
Db      1621 TGGCGGTGAAGTCTTGTGAGAGACGCTCCACCTGACCTCAAGGCAAGTTTCTACAG 1680
Qy      1681 AAGCGAGATCTCTGAAGCAGTACAGCAACCCCAATCCGTCCTCATTTGGTGTGCA 1740
Db      1681 AAGCGAGATCTCTGAAGCAGTACAGCAACCCCAATCCGTCCTCATTTGGTGTGCA 1740
Qy      1741 CCCAGAAAGACCCCATCTACATCGTATGAGAGCTTGTGACGGGGGCGACTTCTGACCT 1800
Db      1741 CCCAGAAAGACCCCATCTACATCGTATGAGAGCTTGTGACGGGGGCGACTTCTGACCT 1800
Qy      1801 TCCTCCGCAACGAGGGGGCCCGCTGCGGTGAAGACTGTGTCAGATGTGGGGGATG 1860
Db      1801 TCCTCCGCAACGAGGGGGCCCGCTGCGGTGAAGACTGTGTCAGATGTGGGGGATG 1860
Qy      1861 CAGCTCTGGCATAGAGTACTTGGAGAGCAAGTGTGATCCACCGGAGACTGTGCTC 1920
Db      1861 CAGCTCTGGCATAGAGTACTTGGAGAGCAAGTGTGATCCACCGGAGACTGTGCTC 1920
Qy      1921 GGAATGCTGTGACAGAGAAATGTCTCTGAAGATCAGTGAATTTGGGATGTCCCGAG 1980
Db      1921 GGAATGCTGTGACAGAGAAATGTCTCTGAAGATCAGTGAATTTGGGATGTCCCGAG 1980
Qy      1981 AGGAAGCCGATGGGGTCTATGACCTCAGGGGGCTCAGACAAGTCCCGTGAAGTGA 2040
Db      1981 AGGAAGCCGATGGGGTCTATGACCTCAGGGGGCTCAGACAAGTCCCGTGAAGTGA 2040
Qy      2041 CCGCACCTGAGGGCCCTTAATAGCGCCGCTACTCTCCGAAAGCGAGTGTGAGCTTTG 2100
Db      2041 CCGCACCTGAGGGCCCTTAATAGCGCCGCTACTCTCCGAAAGCGAGTGTGAGCTTTG 2100
Qy      2101 GCATCTTGTCTGTGGAGACCTTCAAGCCTGAGGGGCTCCCTCTATCCAACTCAGCAATC 2160
Db      2101 GCATCTTGTCTGTGGAGACCTTCAAGCCTGAGGGGCTCCCTCTATCCAACTCAGCAATC 2160
Qy      2161 AGCAGACACGGAGATTGTGAGAAAGGGGGCCGTCTGCTGCCAGAGCTGTCTTG 2220
Db      2161 AGCAGACACGGAGATTGTGAGAAAGGGGGCCGTCTGCTGCCAGAGCTGTCTTG 2220
Qy      2221 ATGCCGTGTTCAAGCTTCATAGAGACATGCTGGGCTTATAGACCTTGGGAGCGCCAGCT 2280
Db      2221 ATGCCGTGTTCAAGCTTCATAGAGACATGCTGGGCTTATAGACCTTGGGAGCGCCAGCT 2280
Qy      2281 TCAGACCACTCTACAGAGCTCAGAGCATCCGAAAGCGGATCGTGAAGCTGGGACC 2340
Db      2281 TCAGACCACTCTACAGAGCTCAGAGCATCCGAAAGCGGATCGTGAAGCTGGGACC 2340
Qy      2341 CCTTCTCAAGCTGTGGCTCTGACAGGCTTAGAGTCACTCTCAGCGGCTCAGCTCA 2400
Db      2341 CCTTCTCAAGCTGTGGCTCTGACAGGCTTAGAGTCACTCTCAGCGGCTCAGCTCA 2400
Qy      2401 TAGTGTGACAGCTCTTCAAGTCTCTGACCTCCGACCAAGCATCCAGCTGGGCAAG 2460
Db      2401 TAGTGTGACAGCTCTTCAAGTCTCTGACCTCCGACCAAGCATCCAGCTGGGCAAG 2460
Qy      2461 ATCAGAGCGCGTGTCTCTGTGTCCCTGTGCTGTGCAAGGGCTTCTCTTCCGGGCAAG 2520
Db      2461 ATCAGAGCGCGTGTCTCTGTGTCCCTGTGCTGTGCAAGGGCTTCTCTTCCGGGCAAG 2520
Qy      2521 AACAAATAAACCACTGTGCCCACTGAAAAA 2580

```

```

Db      2521 AACAAATAAACCACTGTGCCCACTGAAAAA 2580
Qy      2581 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2640
Db      2581 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2640
Qy      2641 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674
Db      2641 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674

```

```

RESULT 3
US-10-240-965-256
; Sequence 256, Application US/10240965
; Publication No. US20030165924A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: SHIFFMAN, DOV
; APPLICANT: SOMOGYI, Roland
; APPLICANT: LAMN, Richard M.
; APPLICANT: SELHAMER, Jeffrey J.
; APPLICANT: PORTER, Gordon J.
; APPLICANT: MIKITH, Thomas
; APPLICANT: TAL, Julie
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
; FILE REFERENCE: PA-0025 PCT
; CURRENT APPLICATION NUMBER: US/10/240,965
; PRIOR FILING DATE: 2002-10-04
; PRIOR FILING DATE: 60/195,106
; NUMBER OF SEQ ID NOS: 276
; SOFTWARE: PERL Program
; SEQ ID NO 256
; LENGTH: 2889
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030165924A1 997347.6
US-10-240-965-256

```

```

Query Match      87.0%; Score 2326.4; DB 16; Length 2889,
Best Local Similarity 92.3%; Pred. No. 0;
Matches 2558; Conservative 0; Mismatches 1; Indels 211; Gaps 2;

Qy      2 CCGGGGTCCGACCCGGGCTGAGTGGTCCGAGCCCTCCAGAGCACTGCCGTGGG 61
Db      118 CCGGGGTCCGACCCGGGCTGAGTGGTCCGAGCCCTCCAGAGCACTGCCGTGGG 177
Qy      62 GAACAGCACTAATGGGCTTCTTCTGAGCTGTGACGCCCAAGGCGCAAGGGTCTTGA 121
Db      178 GAACAGCACTAATGGGCTTCTTCTGAGCTGTGACGCCCAAGGCGCAAGGGTCTTGA 237
Qy      122 GCAATGACGAGAGCCGAGCTTCTCTACTGAGAGGCGATGAGAAATGATGGCCAGCG 181
Db      238 GCAATGACGAGAGCCGAGCTTCTCTACTGAGAGGCGATGAGAAATGATGGCCAGCG 297
Qy      182 GGTCAAGATGACAGGAGATATGACGAGACTGCTTACCAACATGCTCCCTCAGGAGAGTGG 241
Db      298 GGTCAAGATGACAGGAGATATGACGAGACTGCTTACCAACATGCTCCCTCAGGAGAGTGG 357
Qy      242 GGGCCAGAGCCGGGCGATCAGCCCTGACAGCCCAATCACTAGTCTGTGGGCTGAGATCAC 301
Db      358 GGGCCAGAGCCGGGCGATCAGCCCTGACAGCCCAATCACTAGTCTGTGGGCTGAGATCAC 417
Qy      302 CAGCCAACTGA - GGGCTGAGCCCGCTTGTCTGCGGAGAGCAGCAGAGATCTGAATCAG 360
Db      418 CAGCCAACTGAAGGGGCTGAGCCGCTTGTCTGCGGAGAGCAGCAGAGATCTGAATCAG 477
Qy      361 GGGCCCTGAGCAAGCTGAGCTGTCTATCCGGGAAAGGAGAGAGCTTCCCAAGACTTACA 420
Db      478 GGGCCCTGAGCAAGCTGAGCTGTCTATCCGGGAAAGGAGAGAGCTTCCCAAGACTTACA 537

```


OY	421	CCGACGACGTGGCAGACGCTGCAAGCAGGAGCTTCACCAAGACCCACACGCCAGACATTTGAGA	480
Db	538	GCGAGCAGTGGCAGCAGCTGCAAGCAGGAGCTTCACCAAGACCCACACGCCAGACATTTGAGA	597
OY	481	AGCTGAAGCAGCAGTACCGAGCTCTGGCAGCGGAGCAGTGGGCCCAAGGCCAAGGCGCAGTACC	540
Db	598	AGCTGAAGCAGCAGTACCGAGCTCTGGCAGCGGAGCAGTGGGCCCAAGGCCAAGGCGCAGTACC	657
OY	541	AGGAGGCGCAGAAAGACAAAGACCGGTGACAAAGCCCAAGACAAAGTATGTGCGCAGCTGT	600
Db	658	AGGAGGCGCAGAAAGACAAAGACCGGTGACAAAGCCCAAGACAAAGTATGTGCGCAGCTGT	717
OY	601	GGAAAGCTTTTGTGCTCACCAACACCGCTATGTGTGTGGGCTGTGGGCTGTGCGACTACAC	660
Db	718	GGAAAGCTTTTGTGCTCACCAACACCGCTATGTGTGTGGGCTGTGGGCTGTGCGACTACAC	777
OY	661	ACGAGCACACCAACCGAGCTCCGTCGCGCCCGGCGCTGTGCGGTCATCTGACAGGACCTGACAG	720
Db	778	ACGAGCACACCAACCGAGCTCCGTCGCGCCCGGCGCTGTGCGGTCATCTGACAGGACCTGACAG	837
OY	721	AGGAGATGGCTTGCATCTCTGAAGAGATCTTGCAGAAATACCTTGAGATTAGACGCTTGG	780
Db	838	AGGAGATGGCTTGCATCTCTGAAGAGATCTTGCAGAAATACCTTGAGATTAGACGCTTGG	897
OY	781	TGCAGGATGAGGTGTGGCCATTACACCGGAGATGTGCTGACGTGTGCTGCCCATCCAGC	840
Db	898	TGCAGGATGAGGTGTGGCCATTACACCGGAGATGTGCTGACGTGTGCTGCCCATCCAGC	957
OY	841	CTGAGGCTGATATCCAAAGGCTTCCGTGGGACAGTATGGGTCCGACCTGACGTTCCACCT	900
Db	958	CTGAGGCTGATATCCAAAGGCTTCCGTGGGACAGTATGGGTCCGACCTGACGTTCCACCT	1017
OY	901	GTGTACACTTGCATGATGATCACTGTGCTTGAAGAGGGTGAACCCGCTGAGCCTGGGAGCTTC	960
Db	1018	GTGTACACTTGCATGATGATCACTGTGCTTGAAGAGGGTGAACCCGCTGAGCCTGGGAGCTTC	1077
OY	961	AGCTGAACGAGCTGACTGTGTGAAGACGTGCAGCACACGCTGACCTCACTGACAGATGAGC	1020
Db	1078	AGCTGAACGAGCTGACTGTGTGAAGACGTGCAGCACACGCTGACCTCACTGACAGATGAGC	1137
OY	1021	TGGCTGTGGCCACCGAAGTGTGTTCACGGCCGACAGAGATGTGTTACGAGCTGCACACGG	1080
Db	1138	TGGCTGTGGCCACCGAAGTGTGTTCACGGCCGACAGAGATGTGTTACGAGCTGCACACGG	1197
OY	1081	AGCTCCGGAATGGAAGAGGAGAACCCACCCCGGAGCGGGGTGACGTCTGTGGGCAAGA	1140
Db	1198	AGCTCCGGAATGGAAGAGGAGAACCCACCCCGGAGCGGGGTGACGTCTGTGGGCAAGA	1257
OY	1141	GGCAGTGTCTCAGAAAGCACTGTGACGGGGCTGTGACGTATGCGCTGTGCACGAGCCAAAC	1200
Db	1258	GGCAGTGTCTCAGAAAGCACTGTGACGGGGCTGTGACGTATGCGCTGTGCACGAGCCAAAC	1317
OY	1201	TGCAGGCCCCAGCAGAGATTGCTGTGCAGACCAAGCTGTGAGACACTTGGGCCCCGGCGAGCCCC	1260
Db	1318	TGCAGGCCCCAGCAGAGATTGCTGTGCAGACCAAGCTGTGAGACACTTGGGCCCCGGCGAGCCCC	1377
OY	1261	CGCCTGTGCTGCTCTGTGAGATGACGCGCACTCCACGTCGTCCTCGGAGCAGAGAGCGAG	1320
Db	1378	CGCCTGTGCTGCTCTGTGAGATGACGCGCACTCCACGTCGTCCTCGGAGCAGAGAGCGAG	1437
OY	1321	AGGGGGGAGAGACACCCACGCTGTGAGATCTTTAAGAGCCACATCTCAGGAATCTTCCGCC	1380
Db	1438	AGGGGGGAGAGAGACCCACGCTGTGAGATCTTTAAGAGCCACATCTCAGGAATCTTCCGCC	1497
OY	1381	CCAGTTCTTC-----	1391
Db	1498	CCAAGTTCTGCTCCCTCCACCGCTGACCTCATTCGGAAGGTGCAGAAAGCCCTGTGATG	1557
OY	1392	-----	1391
Db	1558	AGCAGCTGTGTAACACGGGGCCATCCCGAAGGACAGAGGTGCTGAGCTGTGTCGACT	1617
OY	1392	-----	1391

Db	1618	CTGGGACCTTCTCTGATGCGGGAAGCAAGGCAAGCAGAGTACGTGTCGCGTCTG	1617
Qy	1392	-----AACCTGTACCAGCTGAAG	1410
Db	1678	GGGATGCTCTCCCGGACCTTCATCATCCAGTCTTGGATTAACCTGAACCGACTGGAA	1737
Qy	1411	GGGAAGCTTTCCTAGCATTCCTTGTGATTCGACACACTATAGACCCAGAGCCCC	1470
Db	1738	GGGAAGCTTTCCTAGCATTCCTTGTGATTCGACACACTATAGACCCAGAGCCCC	1797
Qy	1471	TCACCAAGAAAGATGATGTGTTGCTCTGCAAGAGGCTGTGCCCAAGACAATGAGGTGCTGA	1530
Db	1798	TCACCAAGAAAGATGATGTGTTGCTCTGCAAGAGGCTGTGCCCAAGACAATGAGGTGCTGA	1857
Qy	1511	ACCATAGAGACCTGATGTGTTGGGTAGACAGATTTGACCGGGGACCTTTGGCGAGTGTTC	1590
Db	1858	ACCATAGAGACCTGATGTGTTGGGTAGACAGATTTGACCGGGGACCTTTGGCGAGTGTTC	1917
Qy	1591	GCGAAGCTGTCGAGCCGACCAACCCCTGGTGGGAGTGTGAAGTCTTGTGCAAGACGCTCC	1650
Db	1918	GCGAAGCTGTCGAGCCGACCAACCCCTGGTGGGAGTGTGAAGTCTTGTGCAAGACGCTCC	1977
Qy	1651	CACCTGACCTTGAAAGGCCAAGTTTCTACAGAAAGCAGAGATCTTAAGCAGTACGCCACC	1710
Db	1978	CACCTGACCTTGAAAGGCCAAGTTTCTACAGAAAGCAGAGATCTTAAGCAGTACGCCACC	2037
Qy	1711	CCAAATGTCGCTCTCATGTGTCTGACCCAGAAAGCAGCCCATCTACATGTCATG	2097
Db	2038	CCAAATGTCGCTCTCATGTGTCTGACCCAGAAAGCAGCCCATCTACATGTCATG	2097
Qy	1771	AGCTTGTGCAAGGGGGGACCTTCCGTGACCTTCCCGACGGAAGGGGGCCGCGCGGG	1830
Db	2098	AGCTTGTGCAAGGGGGGACCTTCCGTGACCTTCCCGACGGAAGGGGGCCGCGCGGG	2155
Qy	1831	TGAAGACTCTGCTCAGATGTGTGGGAGATGACGCTGTGAGCATGAGTAACTTGAGAGCA	1890
Db	2158	TGAAGACTCTGCTCAGATGTGTGGGAGATGACGCTGTGAGCATGAGTAACTTGAGAGCA	2217
Qy	1891	AGTGTGATCATCAACCGGAGCCTGTGCTCGGAACTGTCTGTGTACAGAGAAATGTCC	1958
Db	2218	AGTGTGATCATCAACCGGAGCCTGTGCTCGGAACTGTCTGTGTGTACAGAGAAATGTCC	2277
Qy	1951	TGAAGATCAGTGAATTTGGGATGTCCCGAGAGGAAGCCGANTGGGATCTATACAGCTCAG	2010
Db	2278	TGAAGATCAGTGAATTTGGGATGTGTCCGAGAGGAAGCCGANTGGGATCTATACAGCTCAG	2337
Qy	2011	GGGGCTTCAGACAAGTCCCGCTGAAAGTGAACCGACCTGAGGCCCTTAACTACGGCGCT	2070
Db	2338	GGGGCTTCAGACAAGTCCCGCTGAAAGTGAACCGACCTGAGGCCCTTAACTACGGCGCT	2397
Qy	2071	ACTCTCTCCGAAAGGAGCGTGTGAGACTTTGGCATTTGTGCGAAGACTTCAGCTCG	2130
Db	2398	ACTCTCTCCGAAAGGAGCGTGTGAGACTTTGGCATTTGTGCGAAGACTTCAGCTCG	2457
Qy	2131	GGGCGCTCCCGCTATCCCACTCGACATCAAGACACCGGAGTGTGTGAGAGAGGGGG	2190
Db	2458	GGGCGCTCCCGCTATCCCACTCGACATCAAGACACCGGAGTGTGTGAGAGAGGGGG	2517
Qy	2191	GCGCTGTGCTCCGCGCCAGAGCTGTGCTCTGTATGCGGTGTTCAGGCTCAATGAGCAGTCT	2250
Db	2518	GCGCTGTGCTCCGCGCCAGAGCTGTGCTCTGTATGCGGTGTTCAGGCTCAATGAGCAGTCT	2577
Qy	2251	GGGCTATAGACCTTGGGAGCGGCGCCAGCTTCAGACCACTTACAGAGACTTGAGAGCA	2310
Db	2578	GGGCTATAGACCTTGGGAGCGGCGCCAGCTTCAGACCACTTACAGAGACTTGAGAGCA	2637
Qy	2311	TCGGAAGCGGATTCGATGAGGCTGTGGAACCCCTTCTCAAGCTGTGTGCTCTCAGGCG	2370
Db	2638	TCGGAAGCGGATTCGATGAGGCTGTGGAACCCCTTCTCAAGCTGTGTGCTCTCAGGCG	2697
Qy	2371	TAGGTGAGCTTCTTCAAGCGGCTCAAGCTCATATCTGAACAGCTCTTCAAGCTCTGAAGCT	2430

Query Match 12.2%; Score 326.8; DB 10; Length 449;
Best Local Similarity 99.4%; Pred. No. 2.2e-41;
Matches 338; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1778 GGAGGGGGGCGATTCTCTGACCTTCTCCGCAAGAGGGGGCCCGCTCGGGTGAAGAC 1837
DB 120 GGGAGGGGGCGATTCTCTGACCTTCTCCGCAAGAGGGGGCCCGCTCGGGTGAAGAC 179

QY 1838 TCTGCTGCAGATGATGGGGGATGACAGCTGCTGCGATGAGTACTTGAAGACAGTGTCTG 1897
DB 180 TCTGCTGCAGATGATGGGGGATGACAGCTGCTGCGATGAGTACTTGAAGACAGTGTCTG 239

QY 1898 CATCCACCGGGGACCTGGCTGCTCGGAACCTGGTGAAGAGAAAGTCTCTGAAGAT 1957
DB 240 CATCCACCGGGGACCTGGCTGCTCGGAACCTGGTGAAGAGAAAGTCTCTGAAGAT 299

QY 1958 CAGTGACTTTGGGATGCTCCGAGAGAGAACCGGATGGGCTTATGACAGCTTCAAGGGGCGCT 2017
DB 300 CAGTGACTTTGGGATGCTCCGAGAGAGAACCGGATGGGCTTATGACAGCTTCAAGGGGCGCT 359

QY 2018 CAGACAGTCCCGGTGAAGTGAACCGCACCTGAGCCCTTAACAGCGCCGCTACTCTTC 2077
DB 360 CAGACAGTCCCGGTGAAGTGAACCGCACCTGAGCCCTTAACAGCGCCGCTACTCTTC 419

QY 2078 CGAAGCGAGCTGTGAGCTTTGGCATCTT 2107
DB 420 CGAAGCGAGCTGTGAGCTTTGGCATCTT 449

RESULT 7
US-09-948-802-5
; Sequence 5, Application US/09948802
; Publication No. US2002012981A1
; GENERAL INFORMATION:
; APPLICANT: ROBINSON, KEITH E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; FILE REFERENCE: NMI-090
; CURRENT APPLICATION NUMBER: US/09/948,802
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/387,212
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-09-948-802-5

Query Match 11.1%; Score 296.2; DB 9; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.1e-36;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

QY 1684 CGAGATCTCTGAAGCAAGTACAGCCCAACCATCTGCTCATTTGTTCTGACCC 1743
DB 20 CNAGATCTCTGAAGCAAGTACAGCCCAACCATCTGCTCATTTGTTCTGACCC 79

QY 1744 AGAAGAGCCCATCTACATCTGATGAGCTTGTGAGGGGGCGACTTCTGACCTTTC 1803
DB 80 AGAAGAGCCCATCTACATCTGATGAGCTTGTGAGGGGGCGACTTCTGACCTTTC 139

QY 1804 TCCGCAAGAGGGGGCCCGCTGCGGGTGAAGACTCTGCTGAGATGATGGAGATGACG 1863
DB 140 TCCGCAAGAGGGGGCCCGCTGCGGGTGAAGACTCTGCTGAGATGATGGAGATGACG 199

QY 1864 CTGCTGGCATGAGTACTTGAAGAGCAAGTGTGATCCACCGGGAACCTTGCTGCGA 1923
DB 200 CTGCTGGCATGAGTACTTGAAGAGCAAGTGTGATCCACCGGGAACCTTGCTGCGA 259

QY 1924 ACTGCTGTGTGACAGGAAGATGTCTGTAAG-ATCAGTGAATCTT-GGATGTCCGAGA 1981

DB 260 ACTGCTGTGTGACAGGAAGATGTCTGTAAGATCAGATCTTGGGATGTCCGAGA 319

QY 1982 GGAAGCC---GATGGGTCTATGACAGCTCAGGGGGCTTCAG 2020
DB 320 GGAAGCCGATTTGGGGGTCTATGACAGCTCAGGGGGCTTCAG 361

RESULT 8
US-10-121-925-5
; Sequence 5, Application US/10121925
; Publication No. US2003010450A1
; GENERAL INFORMATION:
; APPLICANT: ROBINSON, KEITH E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; FILE REFERENCE: NMI-090
; CURRENT APPLICATION NUMBER: US/10/121,925
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: US/09/948,802
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/387,212
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-10-121-925-5

Query Match 11.1%; Score 296.2; DB 15; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.1e-36;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

QY 1684 CGAGATCTCTGAAGCAAGTACAGCCCAACCATCTGCTCATTTGTTCTGACCC 1743
DB 20 CNAGATCTCTGAAGCAAGTACAGCCCAACCATCTGCTCATTTGTTCTGACCC 79

QY 1744 AGAAGAGCCCATCTACATCTGATGAGCTTGTGAGGGGGCGACTTCTGACCTTTC 1803
DB 80 AGAAGAGCCCATCTACATCTGATGAGCTTGTGAGGGGGCGACTTCTGACCTTTC 139

QY 1804 TCCGCAAGAGGGGGCCCGCTGCGGGTGAAGACTCTGCTGAGATGATGGAGATGACG 1863
DB 140 TCCGCAAGAGGGGGCCCGCTGCGGGTGAAGACTCTGCTGAGATGATGGAGATGACG 199

QY 1864 CTGCTGGCATGAGTACTTGAAGAGCAAGTGTGATCCACCGGGAACCTTGCTGCGA 1923
DB 200 CTGCTGGCATGAGTACTTGAAGAGCAAGTGTGATCCACCGGGAACCTTGCTGCGA 259

QY 1924 ACTGCTGTGTGACAGGAAGATGTCTGTAAG-ATCAGTGAATCTT-GGATGTCCGAGA 1981
DB 260 ACTGCTGTGTGACAGGAAGATGTCTGTAAGATCAGATCTTGGGATGTCCGAGA 319

QY 1982 GGAAGCC---GATGGGTCTATGACAGCTCAGGGGGCTTCAG 2020
DB 320 GGAAGCCGATTTGGGGGTCTATGACAGCTCAGGGGGCTTCAG 361

RESULT 9
US-10-280-576-19
; Sequence 19, Application US/10280576
; Publication No. US20040044405A1
; GENERAL INFORMATION:
; APPLICANT: Wolfe, Matthew R.
; TITLE OF INVENTION: VASCULAR STENT OR GRAFT COATED OR IMPREGNATED WITH PROTEIN
; FILE REFERENCE: 09820.189
; CURRENT APPLICATION NUMBER: US/10/280,576
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 60/343,732

PRIOR FILING DATE: 2001-10-25
 NUMBER OF SEQ ID NOS: 25
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 19
 LENGTH: 1779
 TYPE: DNA
 ORGANISM: Drosophila melanogaster
 US-10-280-576-19

Query Match 10.2%; Score 271.6; DB 17; Length 1779;
 Best Local Similarity 57.0%; Pred. No. 8.5e-33;
 Matches 518; Conservative 0; Mismatches 384; Indels 6; Gaps 1;

1398 TACCGACTGGAAGGGAGGCTTCTTAGCATCTTTGGCTATCGACCACTACTAGAC 1457
 844 TTCGGTTGAGGAGACCACTATTGCGAGCATCAGAGCTATCATGCACTGATCAC 903
 1458 ACCCAGACCCCTCCACCAAGAAAGTGTGTCTTGCACAGGGCTGTGCCAAGAC 1517
 904 TCGGAATTGCGAGTACCGTGAATCGGGAGCATCTCCAGACCCGTTTCCGGAG 963
 1518 AAGTGTGTCTGAACCATGAGACCTGTGTGTGGGTGAGCAGATTGGAGGGGAATT 1577
 964 CGCTGGAGCTGAGCAACGATGATGTGTAATTCTGAGAGAGATTGTGCGGGAACTTT 1023
 1578 GCGGAAGTGTTCGCGGAGCGCTGCGAGCCGACCAACACCCCTGTGCGGTGAAGTCTGT 1637
 1024 GGGGAGTCTACAGGCCCAACTGAAGTCCACCAACTGATGTGCTGTCAAACTGT 1083
 1638 CGAGAGACGCTCCACCTGACTGACCTCAAGGCCAAGTTTCTACAGGAAGCGAGATCTGAAG 1697
 1084 CGAATGACCTGCGCGACGAAAGAGCGTAATTTCTACAGGAAGGCGCATCTCAAG 1143
 1698 CAGTAAGCAACCCCAACATCGTGTCTCATTTGTGTCTGACCCAGAAAGAGCCCATC 1757
 1144 CAATAGGATCATCAATATGTAATGTAATTTGGAATTTGTGAGAAAGAGCCCATC 1203
 1758 TACATGTCATGAGGCTTGTGCAAGGGGGGCACTTCTGACCTTCTCCGACAGAGGG 1817
 1204 ATGATTTGTATGAAATTTGTGTCTCGGTGTGCTTTAACTTTATTCGCAAGAACTCC 1263
 1818 GCCCGCTCGGGGTGAAGACTGTCTGCAAGATGTGGGGATGCACTGTGCGATGAG 1877
 1264 AATGGCTCCACACTGCGCAAAATGGGCAATGAGAGATGCGGCGGAGGATCGCA 1323
 1878 TACCTGAGAGCAAGTGTGCAATCCACCGGAACTGTGCTGTGGAATGCTGTGAGCA 1937
 1324 TATCTGAGTCCAAAACCTGCATTTCACTGCGATCTGCGCGCGGTAAATGTCGTTGAC 1383
 1938 GAGAAGATGTCCTGAAGATCAGTGACTTTGGAGTGTCCGAGAGGAAGCCGATGGGGTTC 1997
 1384 TTGAGACACAGTGTGAAGATCTCCGATTTCCGAATGTCTGCGAGAA-----GAGAA 1437
 1998 TATGACCTCAGAGGGGCTCAGACAAAGTCCCGTGAAGTGAACCGCACTGAGGCCCTT 2057
 1438 TATATGATTTCCGATGGCATGAACAAATPACTGTGAAGTGAACACTCCCGAGGCTTG 1497
 2058 AACTAGGCGCGTACTCTCCGAAAGCAAGTGTGAAGCTTTGGCATCTTGTCTGTGGAG 2117
 1498 AATTTTGGCAAGTACCTTCGTTGTGCGATGTGTGCTTATGCACTACGATGTGGAG 1557
 2118 ACCTTCAAGCTTGGGGGCTCCCTCTATCCCACTCAGCAATCAGAGACAGGAGATT 2177
 1558 ATCTTTCCAAAGGCGACACACCTTACTCCGAGATACCACTCCAGAGCAGAGAGCGC 1617
 2178 GTGAGAGAGGGGCGCTGTGCTGCTGCCAGAGCTGTGTCTGTATCCGTTCAAGCTTC 2237
 1618 ATGATATCGGAGATATGATATGCCAGCGGAGAGAGCGCCGAGGAGATGTATCCGACTG 1677
 2238 ATGAGACAGTGTGGGCTATGAGCTGTGGAGCGGCGCAAGCTTTCAGACCATCTACAG 2297
 1678 ATGCTCAGTGTGGGAGCGGAGCGCAATCCGCAATTCGATTTGATGATCTCAAT 1737

QY 2298 GAGCTGCA 2305
 DB 1738 GTGATGA 1745

RESULT 10
 US-10-101-510-525
 ; Sequence 525, Application US/10101510
 ; Publication No. US20030148295A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WANG, JACKSON
 ; APPLICANT: WANG, YIXIN
 ; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
 ; FILE REFERENCE: 15117.0012
 ; CURRENT APPLICATION NUMBER: US/10/101,510
 ; PRIOR FILING DATE: 2002-03-20
 ; PRIOR APPLICATION NUMBER: 60/276,947
 ; NUMBER OF SEQ ID NOS: 805
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 525
 ; LENGTH: 3875
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-101-510-525

Query Match 7.5%; Score 201.2; DB 15; Length 3875;
 Best Local Similarity 54.9%; Pred. No. 6.6e-22;
 Matches 423; Conservative 0; Mismatches 338; Indels 9; Gaps 1;

1522 GGGTCTGAACCATGAGACCTGTGTGTGGGTGAGCAGATTGACCGGGGAATTGGCG 1581
 1969 GGGAGCTGATCAGGTGGCTGATGTGAGCACTGTCTATGAGAGAGAGATTGGGG 2028
 1582 AAGTGTTCAGCGGAGCGCTGCGAGCC-----GACAAACCTGTGTGGGTGAAT 1632
 2028 AAGTGTATCGAGGAGACCTGAGGCTCCCGACAGAGACTGCAAGACTGTGGCATTAA 2088
 1633 CTGTGAGAGAGCGCTCCACCTGACCTCAAGGCCAAGTTTCTACAGAAAGGAGATCC 1692
 2088 CTTTAAAGACATCTCCAGGTGGCAAGTGTGAACTTCTTTCAGAGGCACTATTA 2148
 1693 TGAAGAGTACAGCAACCCCAACATGTCGTCTCATTTGTGTCTGACCCAGAACAGC 1752
 2149 TGGGCAAGTTTACCAACCGCATATTTCTGATCTGGAAGCGGTGTCACAAAGCAAGC 2208
 1753 CATCTACATGCTCATGAGCTTGTGCAAGGGGGGCACTTCTGACCTTCTCCGACGG 1812
 2209 CGATCATGATCATCAAGAAATTTATGAGAAATGAGCCCTGATGCTTCTGAGGAGC 2268
 1813 AGGGGGCCCGCTGGGGGTGAAGACTGTGCTGAGATGTGGGGATGTCAGCTGTGCA 1872
 2269 GGAAGACAGCTGGTCTCTGGGCACTGATGTGCGCATGCTGCAAGGCATATGATCTGGCA 2328
 1873 TGAAGTACCTGAGAGCAAGTGTGATCTCAACCGGAGCTGTGCTGTGGAATCTGCTGG 1932
 2329 TGAATCACTCAGTATATCAATTAATGTCACCGGAGCTGTGCTGCAAGAACATCTTGG 2388
 1933 TGAAGAGAAAGATGCTCTGAAGATCAAGTACTTTGGATGTGCCAGAGAGAGCCGATG 1992
 2389 TGAATCAAAACCTGTGTGCAAGGTGTGATCTTTGGCTGATCCCTCTGAGATGACT 2448
 1993 GGGTCTATGACCTCAGGGGGCTCAGACAGTCCCGGTGAAGTGAAGGAGCCGACTGAG 2052
 2449 TTGATGACATATCAAGAACCCAGGAGAGAAATCTTATCCGTTGACAGGCCCTGAA 2508
 2053 CCTTAACTACGCGCTACTCTCCGAAAGCAAGTGTGAGCTTTGGCATCTTGTCTCT 2112
 2509 CCATTGGCCATGAGATCTTCAACAAGCAGGAGGTGTGAGCTTTGGGATTTGTATGT 2568
 2113 GGAAGACTTCAAGCTGTGGGGCTCTCCCTTATCCCACTCAGCAATCAGACAGACAGG 2172
 2569 GGAAGTGTCTGAGCTTTGGGACAAAGCTTATGAGGAGATGAGCAATCAGAGGTTATGA 2628

QY 2173 AGTTGTGAGAGAGGGGGGCGCTGCTGCCAGAGCTGTCTCTGTATGCCGTTC 2232
DB 2629 AGAGCATTAAGAGTGGGTACCGGTGCCCCCTCTGTGAAGCTGCGCCCTGTATG 2688
QY 2233 GGCTCATGAGAGAGTGTGGGCTTATGAGCTGTGGGAGCGCCCACTTC 2282
DB 2689 AGCTCATGAAGAAGTGTGGGATATGACCGTCCCGCGGACCACTTC 2738

RESULT 11
US-10-384-339C-1
Sequence 1, Application US/10384339C
Publication No. US20040175703A1
GENERAL INFORMATION:
APPLICANT: Kreutzer, Roland
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING EXPRESSION OF A TARGET GE
FILE REFERENCE: 20200/2002
CURRENT APPLICATION NUMBER: US/10/384,339C
PRIOR FILING DATE: 2003-03-07
PRIOR APPLICATION NUMBER: PCT/EP02/00152
PRIOR FILING DATE: 2002-01-09
PRIOR APPLICATION NUMBER: DE 10100586.5
PRIOR FILING DATE: 2001-01-09
PRIOR APPLICATION NUMBER: DE 10155280.7
PRIOR FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: DE 10158411.3
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: DE 10160151.4
PRIOR FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 173
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 2955
TYPE: DNA
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
TITLE: Bph A1
PATENT DOCUMENT NUMBER: NM00532
US-10-384-339C-1

Query Match 7.4%; Score 198; DB 18; Length 2955;
Best Local Similarity 54.7%; Pred. No. 26-21;
Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

QY 1522 GGGTGTGAACCATGAGACCTGTGTGGGTGAGAGATTGACGCGGAACTTTGGCG 1581
DB 1874 GGGAGCTTATCCAGCGTGGCTGATGTGGAACCTGTCAATAGAGAAAGAGTTGGGG 1933
QY 1582 AAGTGTTCAGCGGACCGCTGCGAGCC-----GACAAACCCCTGTGGCGGTGAAT 1632
DB 1934 AAGTGTATGAGAGGAGCCCTTCAGGCTCCAGCCAGACAGCTGCAAGACTGTGGCCATTAA 1993
QY 1633 CTGTGTGAGAGAGCTCCCACTGACCTTAAGGCGAAAGTTTCAAGAGAGGAGATCC 1692
DB 1994 CCTTAAAGACATATCCAGGTGCGCAGTGTGGAATCTTTCAGAGAGCACTATGA 2053
QY 1693 TGAAGAGTACAGCCACCCCAACATGTGCGTCTCAATTGTTCTGACCCAGAAAGCAG 1752
DB 2054 TGGGCGCAAGTTTGGCCACCGCATATCTGTATGGAAGCGGTGCGCAAAAGCGAAAG 2113
QY 1753 CCATCTACATCGTATGAGAGCTTGTGAGGGGGGCACTTCTGACCTTCTCCGACGG 1812
DB 2114 CCATCTACATCGTATGAGAGATTTATGAGAAATGCAAGCCCTGATGCTTCTGAGGGAGC 2173
QY 1813 AGGGGGCCGGCTGGGGGGGAAAGACTGTGTGAGATGAGGGGGAAATGACGCTGTGGA 1872
DB 2174 GGGAGAGACAGCTGTGCTCCCTGGGCACTGTGAGCAATGCTGAGGGCAATGACTGGGA 2233
QY 1873 TGGAGTACCTGAGAGCAAGTGTGATCCACCGGAGCCTGTGCTGGAATGTGCTGG 1932
DB 2234 TGAATCACTCAAGTATCAACAATTAATGATCCACCGGAGCCTGTGCGCAAGAAACATCTTGG 2293

QY 1933 TGACAGAGAAAGATGCTTGAAGATCAGTACTTTGGATGTCCGAGAGAAAGCCGATG 1992
DB 2294 TGAATCAAAACCTGTGCTGCAAGGTGTGACTTTTGGCTGTACTGCTTCGTGATGACT 2353
QY 1993 GGGTCTATGACCTTCAGGGGGGCTTCAGAAAGTCCCTGGAATGAGACCGACCTGAG 2052
DB 2354 TTGATGGCACATACGAAACCCAGGGAGAAAGATCCCTATCCGTTGGAACAGCCCTGAAG 2413
QY 2053 CCTTACTACAGCGCGCTACTCTCCGAAAGGAGAGTGTGAGCTTTGGCATTTGCTCT 2112
DB 2414 CCATTGCCATCGAATCTTCACACAGCCAGGAGATGTGTGAGCTTTGGGAATGTATGT 2473
QY 2113 GGGAGACCTTCAGCCCGGGGGGCTCCCTATCCCACTCGAATCAGACAGACGAG 2172
DB 2474 GGGAGTGTCTGAGCTTTGGGAAACAGCTTATGAGGAGATGACATTCAGAGCTTATGA 2533
QY 2173 AGTTGTGAGAGAGGGGGGCGCTGCTGCCAGCTGTGTCTGTATGCCGTTC 2232
DB 2534 AGAGCATTAAGAGATGGGTACCGGTTGCCCTCTGTGTGACTGCGCCCTGTATG 2593
QY 2233 GGCTCATGAGAGATGTGGGCTTATGAGCTTGGGAGCGGCCCACTTC 2282
DB 2594 AGCTCATGAAGAAGTGTGGGATATGACCGTCCCGCGGACCACTTC 2643

RESULT 12
US-09-967-768A-144
Sequence 144, Application US/09967768A
Patent No. US20020150877A1
GENERAL INFORMATION:
APPLICANT: Augustus, Meena
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
FILE REFERENCE: 689290-72
CURRENT APPLICATION NUMBER: US/09/967,768A
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US/60/236,109
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,034
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,111
PRIOR FILING DATE: 2000-09-28
NUMBER OF SEQ ID NOS: 325
SOFTWARE: PatentIn version 3.0
SEQ ID NO 144
LENGTH: 3370
TYPE: DNA
ORGANISM: Homo sapiens
US-09-967-768A-144

Query Match 7.4%; Score 198; DB 9; Length 3370;
Best Local Similarity 54.7%; Pred. No. 26-21;
Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

QY 1522 GGGTGTGAACCATGAGACCTGTGTGGGTGAGAGATTGACGCGGAACTTTGGCG 1581
DB 1967 GGGAGCTTATCCAGCGTGGCTGATGTGGAACCTGTCAATAGAGAAAGAGTTGGGG 2026
QY 1582 AAGTGTTCAGCGGACCGCTGCGAGCC-----GACAAACCCCTGTGGCGGTGAAT 1632
DB 2027 AAGTGTATGAGAGGAGCCCTTCAGGCTCCAGCCAGACAGCTGCAAGACTGTGGCCATTAA 2086
QY 1633 CTGTGTGAGAGAGCTCCCACTGACCTTAAGGCGCAAGTTTCAAGAGAGGAGATCC 1692
DB 2087 CCTTAAAGACATATCCAGGTGCGCAGTGTGGAATCTTCTGAGAGGCAACTATGA 2146
QY 1693 TGAAGAGTACAGCCACCCCAACATGTGCGTCTCAATTGTTCTGACCCAGAAAGCAG 1752
DB 2147 TGGGCGCAAGTTTGGCCACCGCATATCTGTATGGAAGCGGTGCGCAAAAGCGAAAG 2206
QY 1753 CCATCTACATCGTATGAGAGCTTGTGCAAGGGGGGCACTTCTGACCTTCTCCGACGG 1812
DB 2207 CCATCTATATCATCAAGAAATTTATGAGAAATGCAAGCCCTGATGCTTCTGAGGGAGC 2266

QY	1813	AGGGGGCCCGCTCCGGGTGAAGACTCTGCTGCAGAGATGATGGGGAGTACACTCTGGCA	1872
Db	2267	GGAGGACACAGCTGTCCTCCCTGGGACAGCTAGTGGCCATCTGCAGGGCATAGCATCTGGCA	2326
QY	1873	TGAGATACCTGGAGAGCAGAGTGTCTGCATCCACCGGAGCTGGCTGCTTCGGAACTGCTGG	1932
Db	2327	TGAACCTACCTCAGATATCACATATTAATGTCCACCGGACCTGGCTGGCCGAAACATCTTGG	2386
QY	1933	TGACAGAGAAATATGTCCTGAAGATCAGTGACTTTGGGATGTCGCCGAGAGGAAACCGATG	1992
Db	2387	TGAATCAAACCTGTGTCTGCAGAGTGTGTGACTTTGGCTTAATCTGCTCCTGGATGACT	2446
QY	1993	GGGCTTATGCAAGCCTCAGGGGGCCTCAGACACAAGTCCCGTGAAGTGAACGACACTGAG	2052
Db	2447	TTGATGGCACATACGAAACCCAGGGAGGAAAGATTCCTATCCGTTGGCAGACCCCTGAG	2506
QY	2053	CCCTTAACTACGGCGCGTACTCTCTCCGAAAGCAGATGTGAGAGCTTTGGCATTTTGGCT	2112
Db	2507	CCATTGGCCATCGGATCTTCAACAAGCAGACGATGTGTGAGGCTTTGGGATTTGTGATGT	2566
QY	2113	GGGAGACCTTCAGCGCTGGGGGCTCCGCCATATCCAACTCAGCAATCAGACAGACACGGG	2172
Db	2567	GGGAGGTGTGAGCTTTGGGGGACAAAGCTTATGGGGAGATAGCATATAGAGAGTTATGA	2626
QY	2173	AGTTGTGGAGAGGGGGGCGCTGTGCCCTGCCAGAGCTGTGTCTTGATGCCGTTC	2232
Db	2627	AGAGCATTTGAGGATGGGTATCCGTTTGCCTCCCTCTGTGAGCTGGCCCTCCCTCTGTATG	2686
QY	2233	GGGCTCATGGAGCAGTGTGGGCTATATAGGCTCTGGGACAGCGGCCCAAGCTTC	2282
Db	2687	AGCTCATGAGAACTGTGGGCAATATGACCTGTGCCCGCGGACACACTTC	2736

RESULT 13
 US-10-354-358-101
 Sequence 101, Application US/10354358
 Publication No. US20030157082A1
 GENERAL INFORMATION:
 APPLICANT: Millennium Pharmaceuticals, Inc
 APPLICANT: Hunter, John Joseph
 APPLICANT: Macbeth, Kyle J.
 APPLICANT: Tsai, Fong-Ying
 APPLICANT: Leeson, Andrea
 APPLICANT: Lightcap, Eric S.
 APPLICANT: Williamsen, Mark
 APPLICANT: Rudolph-Owen, Laura A.
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
 TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
 TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1662, 17667, 9235,
 TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
 TITLE OF INVENTION: 9252, 9189, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
 TITLE OF INVENTION: 9950, 2100, 9288, 64698, 10480, 20893, 33330, 1586, 9943,
 TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099.
 TITLE OF INVENTION: 2150, 26553, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
 FILE REFERENCE: MET02-020P10NMNM1
 CURRENT APPLICATION NUMBER: US/10/354,358
 CURRENT FILING DATE: 2003-01-30
 PRIOR APPLICATION NUMBER: US 60/353,600
 PRIOR FILING DATE: 2002-01-31
 PRIOR APPLICATION NUMBER: US 60/364,517
 PRIOR FILING DATE: 2002-03-15
 PRIOR APPLICATION NUMBER: US 60/371,075
 PRIOR FILING DATE: 2002-04-09
 PRIOR APPLICATION NUMBER: US 60/371,507
 PRIOR FILING DATE: 2002-04-10
 PRIOR APPLICATION NUMBER: US 60/372,984
 PRIOR FILING DATE: 2002-04-16
 PRIOR APPLICATION NUMBER: US 60/374,194
 PRIOR FILING DATE: 2002-04-19
 PRIOR APPLICATION NUMBER: US 60/382,995
 PRIOR FILING DATE: 2002-05-24
 PRIOR APPLICATION NUMBER: US 60/385,023

```

? PRIOR FILING DATE: 2002-05-31
? PRIOR APPLICATION NUMBER: US 60/388,853
? PRIOR FILING DATE: 2002-06-14
? PRIOR APPLICATION NUMBER: US 60/389,395
? PRIOR FILING DATE: 2002-06-17
? Remaining Prior Application data removed - See File Wrapper or PALM.
? NUMBER OF SEQ ID NOS: 122
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 101
? LENGTH: 3370
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (94)...(3048)
? US-10-354-358-101

Query Match      7.4%; Score 198; DB 16; Length 3370;
Best Local Similarity 54.7%; Pred. No. 2e-21;
Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

```

QY	1522	GGGGTCTGGAACCAATGAGAGCACTGGTGTGTGGGTGAGCAAGATTTGACAGGGGGGAACTTTGGCC	1581
Dp	1967	GGGAGCTTAAATCCAGCGGTGGCTAATGGTTGACACTGTCTCATATGAGGAAAGGAGTTTGGGG	2026
QY	1582	AAAGTTTCAAGCGGACCGCTTGCAGAC-----GACAAACACCTGTGGCGGTGAAGT	1632
Dp	2027	AAAGTATTCAGGGAGCCCTCAAGCGTCCCAAGCCAGAGCTGCMAAGACTGTGGCCATTAAAG	2086
QY	1633	CTTGTGAGAGAGGCTCCCACTGACCTCAAGGCGCAAGTTTCTACAGAAAGCGAGATCC	1692
Dp	2087	CCTTAAAAAGACAAATCCCCAGGTGGCCAGTGTGTGAACTTCTCTTCGABAGGCAACTATCA	2146
QY	1693	TGAAGCAGTACAGCCACCCCAACATCGTGGCTCATTTGTGTGTGACCCCAAGACGC	1752
Dp	2147	TGGGCGAGTTTAAGCAACCGCAATATCTGCATCTGAAAGGGGTGTCACAAAGCAAAAGC	2206
QY	1753	CCATCTACATCGTCTATGAGACTGTGCAAGGGGGCGCACTTCTGAACCTTCTCCGCAAGC	1812
Dp	2207	CGATCATGATCATCACAGAAATTTATGAGAAATGAGCGCTCGAATGCCCTTCTTGAGGAGC	2266
QY	1813	AGGGGGCGCGGCTGGCGGTAAAGCTGTGCAGATGTGTGGGGAGATGACGCTGTGGCA	1872
Dp	2267	GGGAGGACCAAGCTGTGTCCTGGGCACTAAGTGGCAATGCTGCAGGGCATTAACATTTGGCA	2326
QY	1873	TGGAGTACCTGAGAGCAAGTGTGTCATCCAGCGGACCTGGCTGTTCGAACTGTGCTGG	1932
Dp	2327	TGAATCTACTCAGTAAATCACAATTATGTCCACGGGACCTGGCTGCCAAGAAACATCTTGG	2386
QY	1933	TGACAGAGAAATATGTCCTGAAGATCAATGATCTTTGGAGATGTCCGAGAGAGACCGATG	1992
Dp	2387	TGAATCAAAACCTGTGCTGCAGAGGTGTCTAATTTGGCTGTGACTCTCGATGATGACT	2446
QY	1993	GGGCTATGAGGCTCAGGGGGGCTCAGACAAAGTCCCGTGAAGTGAACGGCACTGAGG	2052
Dp	2447	TTATGTGCACATACGAAACCCAGGGAGAAAGATCCCTATCCGTTGGACACGCCCTGAAG	2506
QY	2053	CCCTTAATCAACGGCGGCTACTCTCCGCGAAAGCGACGTGTGAGCTTTGGACTTTTGTGCTT	2112
Dp	2507	CCATTTGGCCATCGGATTTTACACACAGCCAGCAATGTGTGAGAGCTTTTGGATTGTGATGT	2566
QY	2113	GGGAGACCTTCAGCTTGGGGGCTCCCTCCCTATATCCAACTTCAGCAATTCAGACAGACGGG	2172
Dp	2567	GGGAGGTGCTAAGCTTTGGGGACAAAGCTTATGGGAGATGAGCAATTCAGAGAGTTATGA	2626
QY	2173	AGTTTGTGAGAAAGGGGGGCGGTCTGCTGCTGCGCCAGAGCTGTGTCTGTGATCCGTGTCA	2233
Dp	2627	AGAGCATTTGAGATGTGGGTACCGGTGTGCCCCCTCTGTGTGACTGCTGCCCCCTCTGTATG	2686
QY	2233	GGCTCATGAGACAGTGTGGGCGCTATATAGCGCTGGGACGGGCGGACGCTTC	2282
Dp	2687	AGCTCATGAAGAACTGCTGGGCAATATACCTGTGCGCCGCGGCAACACTTC	2736

RESULT 14
US-10-210-120-19
Sequence 19, Application US/10210120
Publication No. US2003015736A1
GENERAL INFORMATION:
APPLICANT: Chinnaiyan, Arul M.
APPLICANT: Rubin, Mark A.
APPLICANT: Steekumar, Arun
TITLE OF INVENTION: Expression Profile of Prostate Cancer
FILE REFERENCE: UM-07221
CURRENT APPLICATION NUMBER: US/10/210,120
CURRENT FILING DATE: 2002-08-01
PRIOR APPLICATION NUMBER: US 60/309,581
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: US 60/334,468
PRIOR FILING DATE: 2001-11-15
NUMBER OF SEQ ID NOS: 123
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 19
LENGTH: 3370
TYPE: DNA
ORGANISM: Homo sapiens
US-10-210-120-19

Query Match 7.4%; Score 198; DB 16; Length 3370;
Best Local Similarity 54.7%; Pred. No. 2e-21;
Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

QY 1522 GGGTGTGAAACCATGAGACCTGTGTGGGTGAGAGATTGAGCGGGGAACTTTGGCG 1581
DB 1967 GGGAGCTTGTATCCAGGTGCTATGTGTGACACTGTCTAGAGAGAGAGATTGGGG 2026
QY 1582 AAGTGTTCAGCGGAGCGCTGCGAGCC-----GACAAACCCCTGTGGCGGTGAAGT 1632
DB 2027 AAGTGTATGAGGAGACCTTCAGGCTCCCGAGCAGAGCTGCAAGACTGTGGCATTAAAG 2086
QY 1633 CTGTGTGAGAGAGAGCTTCCCACTGACCTCAAGGCCAAGTTTCTACAGAGAACAGATCC 1692
DB 2087 CCTTAAAGACATATCCCGAGGTGGCCAGTGTGGAACCTTCTTCAAGAGCAACTATCA 2146
QY 1693 TGAAGAGTACAGACCCCAACATCGTGTGCTCAATTGGTGTCTGCAACCAGAGAGC 1752
DB 2147 TGGGCCAGTTTACCCAGCCGCAATTTCTGATCTGAGAGCGTGTGCAAAAGCGAAAGC 2206
QY 1753 CCATCTACATCGTATGAGAGCTTGTGCAAGGGGCGCACTTCTGACCTTCTCCGACGG 1812
DB 2207 CGATCTATGATCATCAGAGATTTATGAGAGATGCAAGCCCTGAGTGCCTTCTGAGGGAGC 2266
QY 1813 AGGGGGCCCGCTGCGGGTGAAGACTGTCTGCAATGTGGGGATGCAAGCTGTGGCA 1872
DB 2267 GGGAGGACCAAGCTGTGCTCCCTGGCAGCTAGTGGCCATGTCTGCAAGGGCATTGGCA 2326
QY 1873 TGAAGTACCTGAGAGCAAGTGTGCAACCGGGGACCTGTGCTCGGAACTGCGCTGG 1932
DB 2327 TGAATCACTCACTCAATTAATGATGTCACCGGAGCTGTGCTGCGAGAAACATTTGG 2386
QY 1933 TGAAGAGAGAGATGTCTGAGAGATCAAGTCTTTGGAGTGTCCGAGAGAGAGCCGATG 1992
DB 2387 TGAATCAAAACCTGTGTGCAAGTGTGTCTGACTTTGGCCGATCTCGCTCTGAGATGCT 2446
QY 1993 GGGTCTATGAGAGCTTCAAGGGGGCTTCAAGACAGTCCCGTGAATGAGCCGACCTGAGG 2052
DB 2447 TTGATGAGCAATTCGAAACCCAGAGAGAGAAAGATCCCTATCCGTTGACAGCCCTGAG 2506
QY 2053 CCCTTAACATCGGCGCTACTCTCCGAAAGAGACGAGTGTGAGCTTTGGATCTTGTCTCT 2112
DB 2507 CCATTTGCCATCGAGATCTTCAACACAGCCAGCAATGTGTGAGCTTTGGATTTGTATGT 2566
QY 2113 GGGAGACCTTTCAGAGCTTGGGGCTCCCGCTATCCCACTCAAGCAATCAAGAGACAGCGG 2172
DB 2567 GGGAGGTGCTGAGCTTTGGGGGCAAGGCTTATGAGGGAGATGACCAATCAAGAGAGTTATGA 2626

QY 2173 AGTTTGTGAGAGAGGGGGGCGGTGCTGCCCTGACCAGAGCGTGTCTGATGCCGTGTCA 2232
DB 2627 AGAGCAATGAGAGTGGATACCGGTTGCCCCCTCTCTGACCTGCCCCCTGTGATG 2686
QY 2233 GGCTCATGAGAGCAAGTGTGGGCTTATGAGCTTGGAGCGGCGCCAGCTTC 2282
DB 2687 AGCTCATGAAAGAACTGTCTGGGCAATGACCGTGTCCCGGGCCAGCACTTC 2736

RESULT 15
US-10-776-827-82
Sequence 82, Application US/10776827
Publication No. US20040132086A1
GENERAL INFORMATION:
APPLICANT: Horwitz, Kathryn
APPLICANT: Richter, Jennifer
TITLE OF INVENTION: Progesterone Receptor-Regulated Gene Expression and Methods Relate
FILE REFERENCE: 2848-39
CURRENT APPLICATION NUMBER: US/10/776,827
CURRENT FILING DATE: 2004-02-10
PRIOR APPLICATION NUMBER: US/09/814,915
PRIOR FILING DATE: 2002-03-21
PRIOR APPLICATION NUMBER: 60/214,870
PRIOR FILING DATE: 2000-06-28
NUMBER OF SEQ ID NOS: 108
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 82
LENGTH: 3370
TYPE: DNA
ORGANISM: Homo sapiens
US-10-776-827-82

Query Match 7.4%; Score 198; DB 18; Length 3370;
Best Local Similarity 54.7%; Pred. No. 2e-21;
Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

QY 1522 GGGTGTGAAACCATGAGACCTGTGTGGGTGAGAGATTGAGCGGGGAACTTTGGCG 1581
DB 1967 GGGAGCTTGTATCCAGGTGCTATGTGTGACACTGTCTAGAGAGAGAGATTGGGG 2026
QY 1582 AAGTGTTCAGCGGAGCGCTGCGAGCC-----GACAAACCCCTGTGGCGGTGAAGT 1632
DB 2027 AAGTGTATGAGGAGACCTTCAGGCTCCCGAGCAGAGCTGCAAGACTGTGGCATTAAAG 2086
QY 1633 CTGTGTGAGAGAGAGCTTCCCACTGACCTCAAGGCCAAGTTTCTACAGAGAACAGATCC 1692
DB 2087 CCTTAAAGACATATCCCGAGGTGGCCAGTGTGGAACCTTCTTGAAGAGCAACTATCA 2146
QY 1693 TGAAGAGTACAGACCCCAACATCGTGTGCTCAATTGGTGTCTGCAACCAGAGAGC 1752
DB 2147 TGGGCCAGTTTACCCAGCCGCAATTTCTGATCTGAGAGGGCCTGTGCAAAAGCGAAAGC 2206
QY 1753 CCATCTACATCGTATGAGAGCTTGTGCAAGGGGCGCACTTCTGACCTTCTCCGACGG 1812
DB 2207 CGATCTATGATCATCAGAGATTTATGAGAGATGCAAGCCCTGAGTGCCTTCTGAGGGAGC 2266
QY 1813 AGGGGGCCCGCTGCGGGTGAAGACTGTCTGCAATGTGGGGATGCAAGCTGTGGCA 1872
DB 2267 GGGAGGACCAAGCTGTGCTCCCTGGCAGCTAGTGGCCATGTCTGCAAGGGCATTAGCACTTGGCA 2326
QY 1873 TGAAGTACCTGAGAGCAAGTGTGCAACCGGGGACCTGTGCTCGGAACTGCGCTGG 1932
DB 2327 TGAATCACTCACTCAATTAATGATGTCACCGGAGCTGTGCTGCGAGAAACATTTGG 2386
QY 1933 TGAAGAGAGAGATGTCTGAGAGATCAAGTCTTTGGAGTGTCCGAGAGAGAGCCGATG 1992
DB 2387 TGAATCAAAACCTGTGTGCAAGTGTGTCTGACTTTGGCCGATCTCGCTCTGAGATGCT 2446
QY 1993 GGGTCTATGAGAGCTTCAAGGGGGCTTCAAGTCCCGTGAATGAGCCGACCTGAGG 2052
DB 2447 TTGATGAGCAATTCGAAACCCAGAGAGAGAAAGATCCCTATCCGTTGAGACAGCCCTGAG 2506


```

Qy 2053 CCCTTAAGGAGGCGGCTACTCCGCGAAGGAGCGTGTGAGCTTTGGCATCTTGTCT 2112
    |||||
Db 2507 CCATTGCCCATGGATCTTCAACAGCCAGCGATGTGTGAGCTTTGGATTGTATGT 2566
    |||||
Qy 2113 GGGAGACTTCAAGCTGGGGGCTCCCTTATCCCAACTCAGCAATCAGCAGACAGGG 2172
    |||||
Db 2567 GGGAGGCTGCTGAGCTTTGGGGACAGCCTTATGGGAGATGAGCAATCAGAGGTTATGA 2626
    |||||
Qy 2173 AGTTGTGAGAGAGGGGGCGGTCTGCGCCGCGAGAGCTGTCTGATGCCGTGTCA 2232
    |||||
Db 2627 AGAGCATTTGAGATGGGTACCGGTTGCCCTCTGTGACTGCCCTTGTGTATG 2686
    |||||
Qy 2233 GGCTCATGGAGCAGTGTCTGGGCTTATGAGCCTGGGAGCGGCCAGCTTC 2282
    |||||
Db 2687 AGCTCATGAAGACTGTGGGATATGACCGTGC CGCGGCCACACTTC 2736
    |||||

```

Search completed: March 19, 2005, 19:55:48
 Job time : 1398.78 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 03:43:36 ; Search time 438.254 Seconds
(without alignments)
9983.718 Million cell updates/sec

Title: US-10-660-763-1

Perfect score: 2674
Sequence: 1 tccgggggtccgcacccgggccc.....aaaaaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	2674	100.0	2674	3	US-09-817-180-1	Sequence 1, Appl1
2	2674	100.0	2674	4	US-10-003-295-1	Sequence 1, Appl1
3	362.4	13.6	15287	3	US-09-817-180-3	Sequence 3, Appl1
4	362.4	13.6	15287	4	US-10-003-295-3	Sequence 3, Appl1
5	296.2	11.1	361	3	US-09-387-212-5	Sequence 5, Appl1
6	296.2	11.1	361	3	US-09-387-212-5	Sequence 5, Appl1
7	223.4	8.4	19152	4	US-09-949-016-12110	Sequence 12110, A
8	223.4	8.4	19152	4	US-09-949-016-15795	Sequence 15795, A
9	198	7.4	3370	4	US-09-814-915A-82	Sequence 82, Appl1
10	185.4	6.9	3921	4	US-09-949-016-628	Sequence 628, Appl1
11	185.4	6.9	3921	4	US-09-949-016-2120	Sequence 2120, Appl1
12	174.6	6.5	3386	4	US-09-949-016-5662	Sequence 5662, Appl1
13	174.6	6.5	3416	2	US-08-357-642A-2	Sequence 2, Appl1
14	174.6	6.5	3416	2	US-08-460-626-2	Sequence 2, Appl1
15	174.6	6.5	3416	4	US-09-016-434-1483	Sequence 1483, Appl1
16	174.6	6.5	4089	4	US-09-300-958A-13	Sequence 13, Appl1
17	170.4	6.4	1611	1	US-07-820-011A-3	Sequence 3, Appl1
18	170.4	6.4	1611	4	US-09-860-473-3	Sequence 3, Appl1
19	170.4	6.4	1611	4	US-09-444-711A-1	Sequence 3, Appl1
20	170.4	6.4	1611	4	US-09-444-711A-1	Sequence 3, Appl1
21	170.4	6.4	1611	5	PCT-US93-00445-3	Sequence 3, Appl1
22	170.4	6.4	2455	4	US-09-949-016-4411	Sequence 4411, Appl1
23	169.4	6.3	3623	4	US-08-306-691B-35	Sequence 35, Appl1
24	169.4	6.3	5434	4	US-09-949-016-927	Sequence 927, Appl1
25	169.4	6.3	5763	4	US-09-949-016-3759	Sequence 3759, Appl1
26	169	6.3	3018	4	US-09-949-016-1097	Sequence 1097, Appl1
27	169	6.3	3018	4	US-09-949-016-4749	Sequence 4749, Appl1

28	169	6.3	4871	4	US-09-799-451-448	Sequence 448, Appl1
29	168.2	6.3	2574	4	US-09-142-529-2	Sequence 2, Appl1
30	168.2	6.3	2574	4	US-10-045-428A-2	Sequence 2, Appl1
31	165.8	6.2	738	2	US-08-604-989A-8	Sequence 8, Appl1
32	165.8	6.2	1398	2	US-08-604-989A-9	Sequence 9, Appl1
33	165.8	6.2	1521	2	US-08-604-989A-10	Sequence 10, Appl1
34	165.8	6.2	1713	3	US-09-741-154-1	Sequence 1, Appl1
35	165.8	6.2	1942	2	US-08-604-989A-11	Sequence 11, Appl1
36	165.8	6.2	2000	3	US-08-426-509A-1	Sequence 1, Appl1
37	165.8	6.2	2000	4	US-08-232-545-1	Sequence 1, Appl1
38	165.8	6.2	2000	5	PCT-US95-05008-1	Sequence 1, Appl1
39	164.4	6.1	1467	4	US-09-579-182-2	Sequence 2, Appl1
40	163.6	6.1	2449	4	US-09-383-630-2	Sequence 1248, Appl1
41	163.6	6.1	5993	3	US-09-383-630-2	Sequence 1, Appl1
42	163.6	6.1	5993	3	US-09-383-630-2	Sequence 2, Appl1
43	162.6	6.1	1548	3	US-09-099-053-1	Sequence 1, Appl1
44	162.2	6.1	2440	1	US-08-160-861-2	Sequence 2, Appl1
45	162.2	6.1	2442	1	US-08-542-363-3	Sequence 3, Appl1

ALIGNMENTS

```
RESULT 1
US-09-817-180-1
; Sequence 1, Application US/09817180
; Patent No. 6340584
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL0001183
; CURRENT APPLICATION NUMBER: US/09/817,180
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2674
; TYPE: DNA
; ORGANISM: Human
US-09-817-180-1

Query Match      100.0%; Score 2674; DB 3; Length 2674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TCCGGGGTCCGACCGGGCTGTGATGTCGACGGCGCTCCAGAGAGAGCTCCCGTGC 60
DB      1 TCCGGGGTCCGACCGGGCTGTGATGTCGACGGCGCTCCAGAGAGAGCTCCCGTGC 60
QY      61 GGAAAGCACTATGGGCTTCTTCTTGAAGCTGTGACGCCCAAGGGCAAGGGGTCTTGC 120
DB      61 GGAAAGCACTATGGGCTTCTTCTTGAAGCTGTGACGCCCAAGGGCAAGGGGTCTTGC 120
QY      121 AGCAATGACAGAGCGGCTTCTTCTTGAAGCTGTGACGCCCAAGGGCAAGGGGTCTTGC 180
DB      121 AGCAATGACAGAGCGGCTTCTTCTTGAAGCTGTGACGCCCAAGGGCAAGGGGTCTTGC 180
QY      181 GGGTCAAGGTGACAGAGAGTATGACGAGCTGTTACCAACATATGCTCCAGAGAGTGTG 240
DB      181 GGGTCAAGGTGACAGAGAGTATGACGAGCTGTTACCAACATATGCTCCAGAGAGTGTG 240
QY      241 GGGGCGGAGCGGGGCGCATCAGCCCTGACAGCCCATCATGATGCTGGGCTGATGATCA 300
DB      241 GGGGCGGAGCGGGGCGCATCAGCCCTGACAGCCCATCATGATGCTGGGCTGATGATCA 300
QY      301 CCAGCCAACTGAGGGGCTGTGAGCGGCTTGTCTGTGGAGAGACCGAGAGATCTGAATCAG 360
DB      301 CCAGCCAACTGAGGGGCTGTGAGCGGCTTGTCTGTGGAGAGACCGAGAGATCTGAATCAG 360
QY      361 GGGCCCTGAGCAAGCTGAGCTGTCTCATCGGGAAAGGAGAGAGCTTCCAGAGACTTACA 420
DB      361 GGGCCCTGAGCAAGCTGAGCTGTCTCATCGGGAAAGGAGAGAGCTTCCAGAGACTTACA 420
```

Db 361 GGCCCTGAGCAAGCTGAGCTGCTCATCCGGGAACGGCAGAGCTTCCGAAGACTTAA 420
QY 421 GGAGAGATGGGAGCAGCTGACAGAGAGCTCAACAAAGCCCAACCCAGGAGATTGAA 480
Db 422 GGAGAGATGGGAGCAGCTGACAGAGAGCTCAACAAAGCCCAACCCAGGAGATTGAA 480
QY 481 AGCTGAAGAGCAGTACCGAGCTCTGGCAAGGAGCAGTGGCCCAAGCCCAAGCCAAATACC 540
Db 481 AGCTGAAGAGCAGTACCGAGCTCTGGCAAGGAGCAGTGGCCCAAGCCCAAGCCAAATACC 540
QY 541 AGGAGCCAGCAAGAACAGAGCCTGTACAAAGCCCAAGCAAGATATGTGCGAGCTGT 600
Db 541 AGGAGCCAGCAAGAACAGAGCCTGTACAAAGCCCAAGCAAGATATGTGCGAGCTGT 600
QY 601 GGAAGCTTTTGTTCACCAACCGCTATATGTGGGCGTGGGGGCTGGCAGCTACAC 660
Db 601 GGAAGCTTTTGTTCACCAACCGCTATATGTGGGCGTGGGGGCTGGCAGCTACAC 660
QY 661 ACCAGACCAACCAACAGCTCTGCTGCGCGGCTGTGCGGCTCACTGACAGGACTGACAG 720
Db 661 ACCAGACCAACCAACAGCTCTGCTGCGCGGCTGTGCGGCTCACTGACAGGACTGACAG 720
QY 721 AGGAGATGGCTTGCATCTTGAAGAGATCTGTGAGGAATACCTGAGATTAGCAGCTGG 780
Db 721 AGGAGATGGCTTGCATCTTGAAGAGATCTGTGAGGAATACCTGAGATTAGCAGCTGG 780
QY 781 TGAAGATAGAGTGTGTGGCCATTCACCGGGAATGTGCTGAGAGTGTGCGCGGATCAGC 840
Db 781 TGAAGATAGAGTGTGTGGCCATTCACCGGGAATGTGCTGAGAGTGTGCGCGGATCAGC 840
QY 841 CTGAGGCTGAGTACCAAGGCTTCTGCGACAGATATGGTTCGCACTGAGCTCCACCT 900
Db 841 CTGAGGCTGAGTACCAAGGCTTCTGCGACAGATATGGTTCGCACTGAGCTCCACCT 900
QY 901 GTGTCACTTTCATGATGATCACTGCTTGAAGAGGATGAAACCGCTGAGGCTGGGAGCTTC 960
Db 901 GTGTCACTTTCATGATGATCACTGCTTGAAGAGGATGAAACCGCTGAGGCTGGGAGCTTC 960
QY 961 AGCTGAACGAGCTGATCTGTGAGAGAGGCTGAGCAACGCTGACCTCAAGTACAGATGAGC 1020
Db 961 AGCTGAACGAGCTGATCTGTGAGAGAGGCTGAGCAACGCTGACCTCAAGTACAGATGAGC 1020
QY 1021 TGGCTGTGCGCAACGAGATGATGTTCAAGCGGAGAGAGATGTTTACGACCTCAACAG 1080
Db 1021 TGGCTGTGCGCAACGAGATGATGTTTCAAGCGGAGAGAGATGTTTACGACCTCAACAG 1080
QY 1081 AGCTCGGAAATGAAGAGAGAACCCACCCCGGAGCGGCTGAGCTGTGGCAAGA 1140
Db 1081 AGCTCGGAAATGAAGAGAGAACCCACCCCGGAGCGGCTGAGCTGTGGCAAGA 1140
QY 1141 GGGAGTGTGCAAGAGCACTGACAGGAGCTGAGAGTGTGAGGCTGTGAGCAGGCAAGC 1200
Db 1141 GGGAGTGTGCAAGAGCACTGACAGGAGCTGAGAGTGTGAGGCTGTGAGCAGGCAAGC 1200
QY 1201 TGAAGGCGCAGAGAGTGTCTGACAGCAAGCTGAGCACTTGGGCGCGGAGCGCC 1260
Db 1201 TGAAGGCGCAGAGAGTGTCTGACAGCAAGCTGAGCACTTGGGCGCGGAGCGCC 1260
QY 1261 CGCTGTGTCTGCTCTGAGAGATGACCGCACTTCACGCTGTCTGTGAGCAGAGAGAG 1320
Db 1261 CGCTGTGTCTGCTCTGAGAGATGACCGCACTTCACGCTGTCTGTGAGCAGAGAGAG 1320
QY 1321 AGGGGGGAGAGAACCCACGCTGAGATCTTAAAGGCCACATCTCAAGAAATCTTCCGC 1380
Db 1321 AGGGGGGAGAGAACCCACGCTGAGATCTTAAAGGCCACATCTCAAGAAATCTTCCGC 1380
QY 1381 CCAAGTTCTCGAAGCTGTACCGAAGGAGAGGAGGCTTCTAGACATCTTGTCTCA 1440
Db 1381 CCAAGTTCTCGAAGCTGTACCGAAGGAGAGGAGGCTTCTAGACATCTTGTCTCA 1440
QY 1441 TGAACCACTTGAAGCAACCAAGCCCTCAACCAAGAGAGTGTGTCTGTGACA 1500
Db 1441 TGAACCACTTGAAGCAACCAAGCCCTCAACCAAGAGAGTGTGTCTGTGACA 1500

QY 1501 GGGCTGTGCGCAAGAGCAAGTGGTGTGAACCATGAGAGCTGTGTGGTGAACAGA 1560
Db 1501 GGGCTGTGCGCAAGAGCAAGTGGTGTGAACCATGAGAGCTGTGTGGTGAACAGA 1560
QY 1561 TTGACGCGGGGAACTTTGGCGAAGTGTTCAGCGGAGCGCTGCGAGCGCAACACCTTG 1620
Db 1561 TTGACGCGGGGAACTTTGGCGAAGTGTTCAGCGGAGCGCTGCGAGCGCAACACCTTG 1620
QY 1621 TGGCGGTGAAGTCTTGTGAGAGAGCTCCCACTGACCTCAAGGCAAGTTCTTACAG 1680
Db 1621 TGGCGGTGAAGTCTTGTGAGAGAGCTCCCACTGACCTCAAGGCAAGTTCTTACAG 1680
QY 1681 AAGCGAGATCTTGAAGCAGTACAGCAACCCCAACCTGCGCTCATTTGGTGTGCA 1740
Db 1681 AAGCGAGATCTTGAAGCAGTACAGCAACCCCAACCTGCGCTCATTTGGTGTGCA 1740
QY 1741 CCCAGAGAGCCCATCTTACATGATGAGCTTGTGAGGAGGAGCGACTTCTGACCT 1800
Db 1741 CCCAGAGAGCCCATCTTACATGATGAGCTTGTGAGGAGGAGCGACTTCTGACCT 1800
QY 1801 TCTTCGCAACGAGAGGAGGCGCGCTGCGGTTGAAGATCTGCTGCAAGTGTGGGAGATG 1860
Db 1801 TCTTCGCAACGAGAGGAGGCGCGCTGCGGTTGAAGATCTGCTGCAAGTGTGGGAGATG 1860
QY 1861 CAGCTGTGCAATGAGTACTTGAAGAGCAAGTGTGACACCGGAGCTGTGCTC 1920
Db 1861 CAGCTGTGCAATGAGTACTTGAAGAGCAAGTGTGACACCGGAGCTGTGCTC 1920
QY 1921 GGAATGCTGTGTGACAGAGAAATGTCTGAGATCAGTATCTTTGGATGTCCGAG 1980
Db 1921 GGAATGCTGTGTGACAGAGAAATGTCTGAGATCAGTATCTTTGGATGTCCGAG 1980
QY 1981 AGGAAGCCGATGGGCTTATGAGCTTCAGGAGGCTTCAGAAAGTCCCGTGAATGGA 2040
Db 1981 AGGAAGCCGATGGGCTTATGAGCTTCAGGAGGCTTCAGAAAGTCCCGTGAATGGA 2040
QY 2041 CCGCACTGAGGCTTAACTACGCGCGCTACTCTCCGAAAGCGAGTGTGAGCTTTG 2100
Db 2041 CCGCACTGAGGCTTAACTACGCGCGCTACTCTCCGAAAGCGAGTGTGAGCTTTG 2100
QY 2101 GCAATCTGTCTGGAAGACCTTCAAGCTGTGGGAGCTTCCCTTATCCCACTCAGCAATC 2160
Db 2101 GCAATCTGTCTGGAAGACCTTCAAGCTGTGGGAGCTTCCCTTATCCCACTCAGCAATC 2160
QY 2161 AGCAGACAGGAGATTTGTGAGAGAGGAGGCGCTGTGCTGCGCAGAGCTGTGCTG 2220
Db 2161 AGCAGACAGGAGATTTGTGAGAGAGGAGGCGCTGTGCTGCGCAGAGCTGTGCTG 2220
QY 2221 ATGCGGTTCAGGCTCATGAGCAGTGTGAGGCTATGAGCTGTGAGCGGCGCAGCT 2280
Db 2221 ATGCGGTTCAGGCTCATGAGCAGTGTGAGGCTATGAGCTGTGAGCGGCGCAGCT 2280
QY 2281 TGAACCATCTTACAGAGAGCTGACAGAGATCCGAAAGCGGCAATCGGTGAGGAGC 2340
Db 2281 TGAACCATCTTACAGAGAGCTGACAGAGATCCGAAAGCGGCAATCGGTGAGGAGC 2340
QY 2341 CCTTCTCAAGCTGTGTGCTGTGAGAGCTTGTGAGCACTCTTCAAGCGGCTTCAAGCTCA 2400
Db 2341 CCTTCTCAAGCTGTGTGCTGTGAGAGCTTGTGAGCACTCTTCAAGCGGCTTCAAGCTCA 2400
QY 2401 TATGCTGAGAGCTCTTCAAGTCTTGAAGCTTGTGAGCACTTGTGAGCACTGCGGAG 2460
Db 2401 TATGCTGAGAGCTCTTCAAGTCTTGAAGCTTGTGAGCACTTGTGAGCACTGCGGAG 2460
QY 2461 ATGACAGCGCGTGTCTCTGTGTCTGTGCTGTGCTGAGGCTTCTTCTGCGGAGAG 2520
Db 2461 ATGACAGCGCGTGTCTCTGTGTCTGTGCTGTGCTGAGGCTTCTTCTGCGGAGAG 2520
QY 2521 AACCAATTAACCACTTGTGCGCACTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
Db 2521 AACCAATTAACCACTTGTGCGCACTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580


```
Db 1681 AACGAGGATCTCGAAGCAGTACAGCACCACCAATCGTCTCATTTGCTGTGCA 1740
Qy 1741 CCCAAGAGCAGCCCATCTACATGTCTAGAGCTTTGCGAGGGGGGAGCTTCTGACT 1800
Db 1741 CCCAAGAGCAGCCCATCTACATGTCTAGAGCTTTGCGAGGGGGGAGCTTCTGACT 1800
Qy 1801 TCCTCGCAGCAGGAGGGGGCCGCTGCGGGGTGAAGACTCTGCTGCAAGATGGTGGGAGT 1860
Db 1801 TCCTCGCAGCAGGAGGGGGCCGCTGCGGGGTGAAGACTCTGCTGCAAGATGGTGGGAGT 1860
Qy 1861 CAGCTCTGCGCATGAGTACCTGGAAGAGCAAGTGTGCTGATCCACCGGAGCTTGCTC 1920
Db 1861 CAGCTCTGCGCATGAGTACCTGGAAGAGCAAGTGTGCTGATCCACCGGAGCTTGCTC 1920
Qy 1921 GGAACCTGCTGTGACAGAGAAATGTCTGAAAGATCACTGATCTTTGGGATGCCGAG 1980
Db 1921 GGAACCTGCTGTGACAGAGAAATGTCTGAAAGATCACTGATCTTTGGGATGCCGAG 1980
Qy 1981 AGGAAGCCGATGAGGATCTATGAGCCCTCAGGGGGCCCTCAGACAAATGCCGTGAGTGA 2040
Db 1981 AGGAAGCCGATGAGGATCTATGAGCCCTCAGGGGGCCCTCAGACAAATGCCGTGAGTGA 2040
Qy 2041 CCGCAGCTGAGGCTTAACTAAGCCGCTACTCTCCGAAAGCGAAGTGTGAGCTTTG 2100
Db 2041 CCGCAGCTGAGGCTTAACTAAGCCGCTACTCTCCGAAAGCGAAGTGTGAGCTTTG 2100
Qy 2101 GCATCTTGTCTGTGAGAGACCTTACGCTGCGGGGCTCCCTCTATCCCACTCAGCAATC 2160
Db 2101 GCATCTTGTCTGTGAGAGACCTTACGCTGCGGGGCTCCCTCTATCCCACTCAGCAATC 2160
Qy 2161 AGAGAAGAGGAGGATTTGTGAGAGAGGGGGGCGCTGCGCTGAGCCAGAGCTGTGCTG 2220
Db 2161 AGAGAAGAGGAGGATTTGTGAGAGAGGGGGGCGCTGCGCTGAGCCAGAGCTGTGCTG 2220
Qy 2221 ATGCGGTGTTCAAGCTCTCAGAGCAGTGTGAGGCTATGAGCCTGAGGCGAGCCAGCT 2280
Db 2221 ATGCGGTGTTCAAGCTCTCAGAGCAGTGTGAGGCTATGAGCCTGAGGCGAGCCAGCT 2280
Qy 2281 TCAGACCATCTACAGAGAGCTGCAAGCATCCGAAAGCGGCACTCGGTGAGGCTGGAGC 2340
Db 2281 TCAGACCATCTACAGAGAGCTGCAAGCATCCGAAAGCGGCACTCGGTGAGGCTGGAGC 2340
Qy 2341 CCTTCTCAAGCTGAGTGGGCTCTGAGAGGCTAGAGTGAAGTCCCTCAGCGGCTCAGCTCA 2400
Db 2341 CCTTCTCAAGCTGAGTGGGCTCTGAGAGGCTAGAGTGAAGTCCCTCAGCGGCTCAGCTCA 2400
Qy 2401 TATGCTGACAGCTCTTCAAGTCTGTGACTCTGCGCAACAGCATCCACTGCGGAGG 2460
Db 2401 TATGCTGACAGCTCTTCAAGTCTGTGACTCTGCGCAACAGCATCCACTGCGGAGG 2460
Qy 2461 ATGAGAGGCGGCTGCTCTCTGTGTGCTGCTGCTGCTGCGCAGAGGCTTCTTCCGGGAGA 2520
Db 2461 ATGAGAGGCGGCTGCTCTCTGTGTGCTGCTGCTGCTGCGCAGAGGCTTCTTCCGGGAGA 2520
Qy 2521 AACCAATTAACCACTTGTGCGCACTGAAAAA 2580
Db 2521 AACCAATTAACCACTTGTGCGCACTGAAAAA 2580
Qy 2581 AAAAAA 2640
Db 2581 AAAAAA 2640
Qy 2641 AAAAAA 2674
Db 2641 AAAAAA 2674
```

RESULT 3
US-09-817-180-3
Sequence 3, Application US/09817180
Patent No. 6340584
GENERAL INFORMATION:

```
APPLICANT: GAN, Weiniu et al.  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: CL001183  
CURRENT APPLICATION NUMBER: US/09/817,180  
CURRENT FILING DATE: 2001-03-27  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 15297  
TYPE: DNA  
ORGANISM: Human  
US-09-817-180-3
```

Query Match 13.6%; Score 362.4; DB 3; Length 15297;
Best Local Similarity 99.7%; Pred. No. 5,1e-55;
Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 2187 GGGGGGCGTGTGAGGCTGCGGAGGCTGTGCTGATGACCGTGTTCAGGCTCATGAGCAG 2246
Db 12938 GGGGGGCGTGTGAGGCTGCGGAGGCTGTGCTGATGACCGTGTTCAGGCTCATGAGCAG 12997
Qy 2247 TGTGTGCTATGAGGCTGCGGAGGCTGCGGAGGCTTACGACCATTTACAGAGAGTGCAG 2306
Db 12998 TGTGTGCTATGAGGCTGCGGAGGCTGCGGAGGCTTACGACCATTTACAGAGAGTGCAG 13057
Qy 2307 AGCATCCGAAAGCGGATGTGAGGCTGCGGAGGCTTCTCAAGCTGTGCTTGA 2366
Db 13058 AGCATCCGAAAGCGGATGTGAGGCTGCGGAGGCTTCTCAAGCTGTGCTTGA 13117
Qy 2367 GGGCTGAGTGCAGGCTCTCAGGAGGCTCCAGCTCATATGCTGACAGCTTTCAAGTCTG 2426
Db 13118 GGGCTGAGTGCAGGCTCTCAGGAGGCTCCAGCTCATATGCTGACAGCTTTCAAGTCTG 13177
Qy 2427 GACTCTGCGACAGCATCAGCTGCGGAGGATGAGCGCGCTGTCTGTGCTC 2486
Db 13178 GACTCTGCGACAGCATCAGCTGCGGAGGATGAGCGCGCTGTCTGTGCTC 13237
Qy 2487 CCTGTGCTGCGAGGCTTCTTCCGGGCAAGAAATTAACCACTTGGCCACTG 2546
Db 13238 CCTGTGCTGCGAGGCTTCTTCCGGGCAAGAAATTAACCACTTGGCCACTG 13297
Qy 2547 AAAA 2550
Db 13298 AAAA 13301
```

RESULT 4
US-10-003-295-3
Sequence 3, Application US/10003295
Patent No. 6686187
GENERAL INFORMATION:
APPLICANT: GAN, Weiniu et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001183DIV
CURRENT APPLICATION NUMBER: US/10/003,295
CURRENT FILING DATE: 2001-12-06
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 15297
TYPE: DNA
ORGANISM: Homo sapiens
US-10-003-295-3

Query Match 13.6%; Score 362.4; DB 4; Length 15297;
Best Local Similarity 99.7%; Pred. No. 5,1e-55;
Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2187 GGGGGGCGTGTGAGGCTGCGGAGGCTGTGCTGATGACCGTGTTCAGGCTCATGAGCAG 2246

DB 12938 GGGGGGCGTCTGCTCCCTGCGCCAGAGCTGTGCTGATGCTGCTTCAAGGCTCATGAGCCAG 12997
QY 2247 TGTGGGCGCTATAGCCTGGGCGAGCGGCCAGCTTACAGACCATCTACAGAGAGCTGCAG 2306
DB 12998 TGTGGGCGCTATAGCCTGGGCGAGCGGCCAGCTTACAGACCATCTACAGAGAGCTGCAG 13057
QY 2307 AGCATCCGAAAGGGGAGCATGCTGAGAGTGGGACCCCTTCTCAAGCTGTGGGCTCTGCA 2366
DB 13058 AGCATCCGAAAGGGGAGCATGCTGAGAGTGGGACCCCTTCTCAAGCTGTGGGCTCTGCA 13117
QY 2367 GGGCTGAGTGAAGCTCTCAGCGGCTCAGACTATATGCTGAGAGCTCTTCAAGCTCTG 2426
DB 13118 GGGCTGAGTGAAGCTCTCAGCGGCTCAGACTATATGCTGAGAGCTCTTCAAGCTCTG 13177
QY 2427 GACTCTGCGCAGCAGCATCCAGACTGCGGCGAGAGTGCAGCGGCTGCTCTGTGTGTC 2486
DB 13178 GACTCTGCGCAGCAGCATCCAGACTGCGGCGAGAGTGCAGCGGCTGCTCTGTGTGTC 13237
QY 2487 CCGTCTGCTGCGCAGGCGCTTCTCTTCCGGGCGAGAAACATTAACCACTTGTGCCACTG 2546
DB 13238 CCGTCTGCTGCGCAGGCGCTTCTCTTCCGGGCGAGAAACATTAACCACTTGTGCCACTG 13297
QY 2547 AAAA 2550
DB 13298 AACCA 13301

RESULT 5
US-09-387-212-5
Sequence 5, Application US/09387212A

Patent No. 6309849
GENERAL INFORMATION:
APPLICANT: ROBINSON, KEITH E.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
FILE REFERENCE: NMT-090
CURRENT APPLICATION NUMBER: US/09/387, 212A
CURRENT FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 5
LENGTH: 361
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-09-387-212-5

Query Match 11.1%; Score 296.2; DB 3; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.1e-43;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

QY 1684 CGAGATCTCGAAGCAGTACAGCCCAACATCGTGTCTCATTTGCTGTGACCC 1743
DB 20 CNAAGATCTGAAAGCAGTACAGCCCAACATCGTGTCTCATTTGCTGTGACCC 79
QY 1744 AGAAGAGCCCATCTATCATGTCATGAGCTTGTGAGGGGGGCACTTCTGACCTTCC 1803
DB 80 AGAAGAGCCCATCTATCATGTCATGAGCTTGTGAGGGGGGCACTTCTGACCTTCC 139
QY 1804 TCCGCAAGAGGGGGGCGCCCTCGCGGTGAAGACTCTGTGCAAGATGTGGGAGATGCAG 1863
DB 140 TCCGCAAGAGGGGGGCGCCCTCGCGGTGAAGACTCTGTGCAAGATGTGGGAGATGCAG 199
QY 1864 CTGCTGGCAGTGAAGTCTTGGAGAGAGTGTGATCCACCGGAGCTTGGGCTGCGGA 1923
DB 200 CTGCTGGCAGTGAAGTCTTGGAGAGAGTGTGATCCACCGGAGCTTGGGCTGCGGA 259
QY 1924 ACTGCTGTGAGAGAGAAATGTCTGTGAAG-ATCAGTACTTT-GGAGTGTCCGAGA 1981
DB 260 ACTGCTGTGAGAGAGAAATGTCTGTGAAGATGTGACTTTGGGAGATGTCCGAGA 319

QY 1982 GGAAGCC---GATGGGCTATGACGCTCAGGGGCGCTCAG 2020
DB 320 GGAAGCCGATTTGGGGCTATGACGCTCAGGGGCGCTCAG 361

RESULT 6
US-09-948-802-5
Sequence 5, Application US/09948802

Patent No. 6465232
GENERAL INFORMATION:
APPLICANT: ROBINSON, KEITH E.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
FILE REFERENCE: NMT-090
CURRENT APPLICATION NUMBER: US/09/948, 802
CURRENT FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: 09/387, 212
PRIOR FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 5
LENGTH: 361
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-09-948-802-5

Query Match 11.1%; Score 296.2; DB 3; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.1e-43;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

QY 1684 CGAGATCTCGAAGCAGTACAGCCCAACATCGTGTCTCATTTGCTGTGACCC 1743
DB 20 CNAAGATCTGAAAGCAGTACAGCCCAACATCGTGTCTCATTTGCTGTGACCC 79
QY 1744 AGAAGAGCCCATCTATCATGTCATGAGCTTGTGAGGGGGGCACTTCTGACCTTCC 1803
DB 80 AGAAGAGCCCATCTATCATGTCATGAGCTTGTGAGGGGGGCACTTCTGACCTTCC 139
QY 1804 TCCGCAAGAGGGGGGCGCCCTCGCGGTGAAGACTCTGTGCAAGATGTGGGAGATGCAG 1863
DB 140 TCCGCAAGAGGGGGGCGCCCTCGCGGTGAAGACTCTGTGCAAGATGTGGGAGATGCAG 199
QY 1864 CTGCTGGCAGTGAAGTCTTGGAGAGAGTGTGATCCACCGGAGCTTGGGCTGCGGA 1923
DB 200 CTGCTGGCAGTGAAGTCTTGGAGAGAGTGTGATCCACCGGAGCTTGGGCTGCGGA 259
QY 1924 ACTGCTGTGAGAGAGAAATGTCTGTGAAG-ATCAGTACTTT-GGAGTGTCCGAGA 1981
DB 260 ACTGCTGTGAGAGAGAAATGTCTGTGAAGATGTGACTTTGGGAGATGTCCGAGA 319
QY 1982 GGAAGCC---GATGGGCTATGACGCTCAGGGGCGCTCAG 2020
DB 320 GGAAGCCGATTTGGGGCTATGACGCTCAGGGGCGCTCAG 361

RESULT 7
US-09-949-016-12110
Sequence 12110, Application US/09949016

Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241, 755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237, 768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231, 498

PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12110
LENGTH: 19152
TYPE: DNA
ORGANISM: Human
US-09-949-016-12110

Query Match
Best Local Similarity 99.6%; Score 223.4; DB 4; Length 19152;
Pred. No. 2,4e-30;
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 60 CGGAACGACACTATGGCTTCTCTCTGAGCTGTGACGCCCCGAGGGCCACGGGGTCTTG 119
DB 18729 CAGAAACGACACTATGGCTTCTCTCTGAGCTGTGACGCCCCGAGGGCCACGGGGTCTTG 18788
QY 120 CAGCAATATGAGAGGCGGAGCTTCTCTGAGCTGTGACGCCCCGAGGGCCACGGGGTCTTG 179
DB 18789 CAGCAATATGAGAGGCGGAGCTTCTCTGAGCTGTGACGCCCCGAGGGCCACGGGGTCTTG 18848
QY 180 CGGGTCAAGAGTGAACAGGAGTATGACAGCTGCTTCAACCATGTCCTTCAGAGACAGT 239
DB 18849 CGGGTCAAGAGTGAACAGGAGTATGACAGCTGCTTCAACCATGTCCTTCAGAGACAGT 18908
QY 240 GGGGGCCAGAGCCGGGCTCATGAGCCCTGACAGCCCCCATCAGTCA 284
DB 18909 GGGGGCCAGAGCCGGGCTCATGAGCCCTGACAGCCCCCATCAGTCA 18953

RESULT 8

US-09-949-016-15795
Sequence 15795, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15795
LENGTH: 19153
TYPE: DNA
ORGANISM: Human
US-09-949-016-15795

Query Match
Best Local Similarity 99.6%; Score 223.4; DB 4; Length 19153;
Pred. No. 2,4e-30;
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 60 CGGAACGACACTATGGCTTCTCTCTGAGCTGTGACGCCCCGAGGGCCACGGGGTCTTG 119
DB 18729 CAGAAACGACACTATGGCTTCTCTCTGAGCTGTGACGCCCCGAGGGCCACGGGGTCTTG 18788
QY 120 CAGCAATATGAGAGGCGGAGCTTCTCTGAGCTGTGACGCCCCGAGGGCCACGGGGTCTTG 179
DB 18789 CAGCAATATGAGAGGCGGAGCTTCTCTGAGCTGTGACGCCCCGAGGGCCACGGGGTCTTG 18848
QY 180 CGGGTCAAGAGTGAACAGGAGTATGACAGCTGCTTCAACCATGTCCTTCAGAGACAGT 239
DB 18849 CGGGTCAAGAGTGAACAGGAGTATGACAGCTGCTTCAACCATGTCCTTCAGAGACAGT 18908
QY 240 GGGGGCCAGAGCCGGGCTCATGAGCCCTGACAGCCCCCATCAGTCA 284
DB 18909 GGGGGCCAGAGCCGGGCTCATGAGCCCTGACAGCCCCCATCAGTCA 18953

DB 18909 GGGGGCCAGAGCCGGGCTCATGAGCCCTGACAGCCCCCATCAGTCA 18953

RESULT 9

US-09-814-915A-82
Sequence 82, Application US/09814915A
Patent No. 6750015
GENERAL INFORMATION:
APPLICANT: Horwitz, Kathryn
TITLE OF INVENTION: Progesterone Receptor-Regulated Gene Expression and Methods Related
THERE TO
FILE REFERENCE: 2848-39
CURRENT APPLICATION NUMBER: US/09/814,915A
CURRENT FILING DATE: 2002-03-21
PRIOR APPLICATION NUMBER: 60/214,870
PRIOR FILING DATE: 2000-06-28
NUMBER OF SEQ ID NOS: 108
SOFTWARE: PatentIn version 3.1
SEQ ID NO 82
LENGTH: 3370
TYPE: DNA
ORGANISM: Homo sapiens
US-09-814-915A-82

Query Match
Best Local Similarity 7.4%; Score 198; DB 4; Length 3370;
Pred. No. 5e-26;
Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

QY 1522 GGGTCTGACACATGAGAGAGCTGTGTTGGTGTGAGCAGATTGACGGGGAACTTTGGCG 1581
DB 1967 GGGAGCTGTATCAGAGCTGTGATGTGTGACACTGTATAGAGAAAGAGTTGGGG 2026
QY 1582 AAGTGTTCAGAGAGCGGCTTCGAGCC-----GACAAACCCCTGTGTGGGTGAAGT 1632
DB 2027 AAGTGTATGAGAGAGCTTCAGAGCTTCAGAGCCAGAGCTGTGAGACTGTGTGATTAAG 2086
QY 1633 CTGTGTGAGAGAGAGCTTCAGAGCTTCAGAGCCAGAGCTGTGAGACTGTGTGATTAAG 1692
DB 2087 CCTTAABAAACATATCCAGAGTGTGTGAGAGCTTCCTGTGAGAGGCAATATCA 2146
QY 1693 TGAAGCAGTACAGAGCCAGAGCTTCAGAGCTTCAGAGCCAGAGCTGTGTGATTAAG 1752
DB 2147 TGGGCGAGTTTACAGAGCCAGAGCTTCAGAGCTTCAGAGCCAGAGCTGTGTGATTAAG 2206
QY 1753 CCATCTACATGATGATGAGAGCTTCAGAGCTTCAGAGCCAGAGCTTCAGAGCTTCAGAG 1812
DB 2207 CCATCTACATGATGATGAGAGCTTCAGAGCTTCAGAGCCAGAGCTTCAGAGCTTCAGAG 2266
QY 1813 AGGGGCGGCGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAG 1872
DB 2267 GGGAGGACAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAG 2326
QY 1873 TGAAGTACCTGAGAGAGAGTGTGTGATTCAGAGCTTCAGAGCTTCAGAGCTTCAGAG 1932
DB 2327 TGAAGTACCTGAGAGAGAGTGTGTGATTCAGAGCTTCAGAGCTTCAGAGCTTCAGAG 2386
QY 1933 TGAAGTACCTGAGAGAGAGTGTGTGATTCAGAGCTTCAGAGCTTCAGAGCTTCAGAG 1992
DB 2387 TGAAGTACCTGAGAGAGAGTGTGTGATTCAGAGCTTCAGAGCTTCAGAGCTTCAGAG 2446
QY 1993 GGGTGTATGAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAG 2052
DB 2447 TGAAGTACCTGAGAGAGAGTGTGTGATTCAGAGCTTCAGAGCTTCAGAGCTTCAGAG 2506
QY 2053 CCCTTAACAGAGCGGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAG 2112
DB 2507 CCATTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAG 2566
QY 2113 GGGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAG 2172
DB 2567 GGGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAG 2626

DB 2457 CGCTGACCGCCCGGAGGACCATTTCTCAACCGAAGTTCACTTCGACGACGATGTGG 2516
QY 2094 AGCTTTGGCATCTGTCTCTGGGAGACCTTCAGCTTGGGGGCTCCCTTATCCCACTTC 2153
DB 2517 AGCTTTGGCATCTGTCTCTGGGAGACCTTCAGCTTGGGGGCTCCCTTATCCCACTTC 2576
QY 2154 AGCAATCAGACAGACGAGGAGTTTGTGAGAAAGGGGGCCGTCTGCTCCGACCAAGCTG 2213
DB 2577 TCCACCAAGAGGTATGAAAGCAATCATATGATGCTTCGCGTCCCAACCCATGAGAC 2636
QY 2214 TGTCTGATGCTGCTTTCAGGCTCATGAGACAGTGTGGGCTTATGAGCTTGGGACGG 2273
DB 2637 TGGCTCTCGGCATCTACCAAGCTCATGATGAGTGTCTGGCAGACGAGCGCTCCGCGC 2696
QY 2274 CCCAGCTTACGACCATCTTACCAAGAGCTGCAAGATC 2312
DB 2697 CCCAGTCTGCTGACATCTGACAGATCTGCAAGACTC 2735

RESULT 12
US-09-949-016-5662
Sequence 5662, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 5662
LENGTH: 3386
TYPE: DNA
ORGANISM: Human
US-09-949-016-5662

Query Match 6.5%; Score 174.6; DB 4; Length 3386;
Best Local Similarity 54.7%; Pred. No. 7.1e-22;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;

QY 1569 GGGAACTTTGGCGAAGTGTTCAGCGGACGCTCGAGCGCAACAACCTGTGGCGGTG 1628
DB 1386 GGGGAGGTCTATGAGGTGTCTACAAATCACAAGGGGAGAAATCATATGATGCTTC 1445
QY 1629 AAGCTTGTGAGAGACGCTCCCACTGACCTCAAGGCCAAGTTTCTACAGAAAGGAG 1688
DB 1446 AAGACTGCAAGAAAGACTGCACTTGACAAACAAGAGAGTTCAATGACGAGGAGTG 1505
QY 1689 ATCTTAAGAGTACAGCAACCCCAACATGTCCTGCTCATATGATGTCTGACACCGAGA 1748
DB 1506 ATATTAAGAACTTCAACACCGCACATCTGTAAGCTGATGCGCAT---CATTTAAGAG 1562
QY 1749 CAGCCCATCTACATCTGATGAGCTTGTGCAAGGGGGGCACTTCTGACCTTCTCCG 1808
DB 1563 GAGCCCACTGATCTATCATGAAATTTATCTCTATGAGGAGCTGGGCCATCTCGGAG 1622
QY 1809 ACGGAGGGGCGCGCTGCGGTGAAGACTCTGCTGCAATGTGTGGGGATGCAAGCTGCT 1868
DB 1623 CGAACAAGAACTTCCGTAAGGTGTCAACCTGCTGTGATCACTGACGATATCAAA 1682
QY 1869 GCGATGAGTACCTGAGAGCAAGTGTGATCCACCGGAGACTGTGCTGCGAACTGC 1928
DB 1683 GCGATGCGCTTACCTGAGAGCAATCACTGCTGCAACGAGGACATTTCTGTCGGAACATC 1742
QY 1929 CTGCTGACAGAGAAATGTCTGAAGATCAGTGAATTTGGATGTCTCCGAGAGAGCC 1988

DB 1743 CTGTGTGCTCTCCCTGAGTGTGTGAAGCTGGGGGACTTTGTCTTCCCGGTACATTTAG 1802
QY 1988 GATGGGCTTATGACGCTCAGGGGGCTTCAGACAAGTCCCCGTGAAGTGAACCGACCT 2048
DB 1803 GACGAGGACTAT---TACAAAGCCCTGTGACTGTCTTCCCATCAAAATGATGTCCCA 1859
QY 2049 GAGGCCCTTAATACAGGCGGCTACTCTCCGAAGGAGAGTGTGAGCTTTGGCATCTTG 2108
DB 1860 GAGTCATTAATCTTCGACGCTTCACAGACGACATGATGTCTGATGTTCCCGGTGTC 1919
QY 2109 CTGTGGAGACCTTCAGCTGGGGGCTTCCCTTATCCCACTGACAAATACAGACACA 2168
DB 1920 ATGTGGAGATCTGAGCTTTGGGAAGCAGCCCTTCTGTGCTGGAGAAACAAGATGTC 1979
QY 2169 CGGAGCTTTGTGAGAAAGGGGGCCGTCTGCTCCCTGCGCAGACTGTCTGATGCGCTG 2228
DB 1980 ATCGGGGTGTGGAAAGAGACCGGCTGCCAAGCTGATCTGTCCACGGGTCTT 2039
QY 2229 TTCAGGCTCATGAGACAGTGTGGGCTTATGAGCTTGGGACGCGGCCAGCTTCAAC 2285
DB 2040 TATACCTCATATACCGCTGCTGTGGACTACGACCCCAATGACCGGCCGCTTCAAC 2096

RESULT 13
US-08-357-642A-2
Sequence 2, Application US/08357642A
Patent No. 5837524
GENERAL INFORMATION:
APPLICANT: Sima Lev
APPLICANT: Joseph Schlessinger
TITLE OF INVENTION: PYK2 RELATED PRODUCTS
TITLE OF INVENTION: AND METHODS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 613 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/357,642A
FILING DATE: December 15, 1994
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
FILING DATE:
NAME: Marburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 209/070
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3416
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic
US-08-357-642A-2

Query Match 6.5%; Score 174.6; DB 2; Length 3416;

Best Local Similarity 54.7%; Pred. No. 7.1e-22;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;

1569 GGGAACTTTGGCGAAGTGTTCACGCGAGCCGCGAACAACCCCTGGTGGCGGTG 1628
1416 GGGGAGGTCTATGAGAGGTGTCTACACAATTCACAAAGGGGAGAAATTCATGATGCTGTC 1475
1629 AAGCTTTGTGAGAGACGCTCCCACTGACCTTAAGGCCAAGTTTCTACAGAGGAGAG 1688
1476 AAGACCTGGAAGAAAGACTGCACTCTGAGCAACAAAGAGAAAGTTGATGAGCGGAGAGT 1535
1689 ATCTGAGAGAGTACAGCAAGCCCAACATGTCGCTGCTCATTTGATGTCGACCCAGAG 1748
1536 ATCATAGAGAACTTCACACACCCGCACTGTAAGTGAATCCGCAAT--CATTTAAGAG 1592
1749 CAGCCCATCTACATCTGATGAGCTTGTGACAGGGGGCGACTTCTGACCTTCTCCG 1808
1593 GAGCCACCTGATCATGATGAAATTTATCCCTATGGGAGCTGGGCCACTGAGAG 1652
1809 ACGGAGGGGGCCGCTGCGGGTGAAGACTCTGCTGCAAGTGGTGGGGAGTGCAGCTGCT 1868
1653 CGGAAACAAGAACTCCCTGAAGGTGCTCACCTGCTGTACTCACTGCAGATATGCATA 1712
1869 GGCATGAGTACCTGAGAGCAAGTGTGATCCACCGGAGCTGTGCTGCGGAATGCTG 1928
1713 GCCATGACCTTACCTGAGAGCAATCACTGCTGCAACAGGACATGCTGCTCCGGAACATC 1772
1929 CTGTGACAGAGAAAGATGTCTGAAGATCACTGATTTGGATGTCTCCGAGAGAGGCC 1988
1773 CTGTGAGCTCCCTGAGTGTGAGAGCTGGGGAATTTGTGTTTCCGGTACATTTAG 1832
1989 GATGGGTCTATGACCTTCAGGGGGCTTCAGAACAGTCCCGTGAAGTGAACCGACCT 2048
1833 GACGAGAGACTAT--TACAAAGCCTGTGATGATGCTCTCCCATCAAAAGATGATGCCA 1889
2049 GAGGACCTTAACACGCGGCTACTCTCCGAAGGAGAGTGTGAGCTTTGGCATCTTG 2108
1890 GAGTCCATTAACCTTCGAGCGCTTCACAGACAGCAAGTGAAGTGTGAGTGTGCGGAGTGC 1949
2109 CTCTGAGAGACCTTCAGCCTGAGGGGCTCCCTATCCCAACTCAGCATCAGACAGCA 2168
1950 ATGTGGAGATCTGAGCTTTGGGAAGCAGCCCTTCTCTGAGTGAAGAACAGATGTC 2009
2169 CGGAGCTTTGTGAGAAAGGGGGCTCTGCTGCTGCGCCAGAGCTGTGTCTGATGCCGTG 2228
2010 ATCGGGGTGTGAGAAAGAGACCGGCTGCCAAGCTGATCTGTCCACCGGTCTT 2069
2229 TTCAGGCTCATGAGAGAGTGGGCTTATGAGCTGGGAGAGGCCCAAGCTTCAGC 2285
2070 TATACCTCATGACCGCTGTGCTGAGACTGACCCCAAGTGAACCGGCCCGCTTACCC 2126

RESULT 14
US-08-460-626-2

Sequence 2, Application US/08460626
Patent No. 5837815
GENERAL INFORMATION:
APPLICANT: SIMA LEV
APPLICANT: JOSEPH SCHLESSINGER
TITLE OF INVENTION: PYK-2 RELATED PRODUCTS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: Storage

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,626
FILING DATE: June 2, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/357,642
FILING DATE: December 15, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 211/121
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3416
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-460-626-2

Query Match 6.5%; Score 174.6; DB 2; Length 3416;
Best Local Similarity 54.7%; Pred. No. 7.1e-22;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;

1569 GGGAACTTTGGCGAAGTGTTCACGCGAGCCGCGAACAACCCCTGGTGGCGGTG 1628
1416 GGGGAGGTCTATGAGAGGTGTCTACACAATTCACAAAGGGGAGAAATTCATGATGCTGTC 1475
1629 AAGCTTTGTGAGAGAGCTCCCACTGACCTCAAGGCCAAGTTTCTACAGAGAGAGG 1688
1476 AAGACCTGGAAGAAAGACTGCACTCTGAGCAACAAAGAGAAAGTTGATGAGCGGAGAGT 1535
1689 ATCTGAGAGACGTACAGCAAGCCCAACATGTCGCTCATTTGATGTCGACCCAGAG 1748
1536 ATCATAGAGAACTTCGACACCCGCACTGTAAGTGAAGTGAAGTGTGAGTGTGCGGAGTGC 1592
1749 CAGCCCATCTACATCTGATGAGCTTTGGAGAGGGGGGAGACTTCTGACCTTCTCCG 1808
1593 GAGCCCACTGTGATCAATGAAATTTATCCCTATGAGGAGAGTGGGCACTACCTGAG 1652
1809 ACGGAGGGGGCCGCTGAGGGTGAAGACTTGTCTGACATGATGATGAGGAGTGCAGCTGCT 1868
1653 CGGAAACAAGAACTCCCTGAAGGTGTCTACCTGTGTGTATCTGACATGAGATATGCCAA 1712
1869 GGCATGAGTACCTGAGAGCAAGTGTGATCCACCGGAGCTGTGCTGCGGAATGCTG 1928
1713 GCCATGACCTTACCTGAGAGCAATCACTGCTGTCAGACAGGAGACATGCTGTCGGAACATC 1772
1929 CTGTGACAGAGAAAGATGTCTGAAGATCACTGATTTGGAGATGTCGCCAGAGAGGCC 1988
1773 CTGTGAGCTCCCTGAGTGTGAGAGCTGGGGAGCTTTGTGCTTTCCTCCGATCATTTAG 1832
1989 GATGGGTCTATGAGAGACTTCAGGGGGCTTCAGCAAGTCCCGGTGAAGTGAACCGGACT 2048
1833 GACGAGAGACTAT--TACAAAGCCTTGTGATCTGCTCTCCCATCAAAAGATGATGCCCA 1889
2049 GAGGACCTTAACACGCGGCTACTCTCCGAAGGAGAGAGTGTGAGCTTTGGCATCTTG 2108
1890 GAGTCCATTAACCTTCGAGCGCTTCACAGACAGCAAGTGAAGTGTGAGTGTGCGGAGTGC 1949
2109 CTCTGAGAGACCTTCAGCTGAGGGGCTCCCTTATCCCAACTCAGCATCAGAGAGCA 2168
1950 ATGTGGAGATCTCTGAGCTTTGGGAAGCAGCCCTTCTTGTGCTGAGAAACAAGATGTC 2009
2169 CGGAGTTTGTGAGAAAGGGGGCTGTGCTGCTGCGCCAGAGCTGTGCTTGAATGCCGTG 2228
2010 ATCGGGGTGTGAGAAAGAGAACCGGCTGTGCCAAGCTGATCTGTCTCACCAGCTT 2069

QY 2229 TTCAGGCTCATGAGAGAGTGTGAGCTTGTGAGCCTTGGGAGAGGCCCAAGCTTACG 2285
DB 2070 TATACCTCATGACCCGCTGTGGAGCTTACGACCCCAAGTGAAGCGGCCCGCTTACCC 2126

RESULT 15
US-09-016-434-1483
Sequence 1483, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 845-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1483:
SEQUENCE CHARACTERISTICS:
LENGTH: 3416 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g988304
US-09-016-434-1483

Query Match 6.5%; Score 174.6; DB 4; Length 3416;
Best Local Similarity 54.7%; Pred. No. 7.1e-22;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;

QY 1569 GGGAACTTTGGCGAAGTGTTCAGCGGAGCGCTTCCGAGCCGCAACAACCTGTGGCGGTG 1628
DB 1416 GGGGAGGTCTATGAGAGGTGTCTACACAAATCACAAAGGGAGAAATCAATGTAGCTGTC 1475
QY 1629 AAGCTTTGTGGAAGAGCGCTCCCACTGAGCCTGAAGGCCAAGTTTCTACAGGAAGGAGG 1688
DB 1476 AAGACCTGCAAGAAAGCTGCACTTGGACAAACAAGAGAAATTCAATGAGCGGAGGTG 1535
QY 1689 ATCTGAAGCAGTACAGCCCAACCCCAATCGTCCGTCTCATTTGATGTGCAACCCAGAAG 1748
DB 1536 ATCATGAAGAACTCGACCAACCCGCAATGTGAAGTGAATGGCAT---CATGAAGAG 1592
QY 1749 CAGCCCATCTACATCGTCATGAGAGCTTGTGCAAGGGGGCGACTTCTGACCTTCTCCGC 1808
DB 1593 GAGCCCACTGATCATGATGAAATGTATCCTATGGGAGAGCTGGGCGCACTACCTGAG 1652

QY 1809 ACGAGGGGGCCCGCTCGCGGTGAAGAATCTGTGTCAGATGATGGGGATGCAAGCTGTG 1868
DB 1653 CGGAACAAGAAATCTCCCTGAAGGTCTCAACCCCTGTGTCTACTCACTGAGATATGCAAA 1712
QY 1869 GGCATGAGATACCTGGAAGAGCAAGTGTGATTCACCGGAGCCTGTGCTGGAACTGC 1928
DB 1713 GCCATGCGCTTACCTGAGAGCATCAACGTGCGACAGGGACATTTGCTCGGAAATATC 1772
QY 1929 CTGTGACAGAAAGAAATCTCTGAAGATCACTGACTTGGGATGTCCGAGAGGAAGCC 1988
DB 1773 CTGTGGCTCTCCCTGAGTGTGAAAGCTGGGGAACTTGTGCTTCCGGTACATTGAG 1832
QY 1989 GATGGGCTTATGAGAGCTCAGAGGGGCTCAGACAAGTCCCGTGAATGSAACCGCACT 2048
DB 1833 GACGAGACTAT---TACAAAGCTCTGTGACTGTCTCCCATCAAAAGATGTCCCA 1889
QY 2049 GAGGCCCTTAAGTACAGCGCGCTACTCTCCGAAAGCAGTGTGAGCTTTGGCATCTTG 2108
DB 1890 GAGTCATTAATCTCCGAGCTTCAAGACAGCAAGTGAAGTGTGATGTTCCCGGTGTC 1949
QY 2109 CTGTGGAGACCTTCAAGCCTGGGGGCTCCCTATCCCAACTCAGCAATGAGCAGACA 2168
DB 1950 ATGTGGAGATCTGAGCTTTGGGAAAGCAGCCCTTCTTGTGCTGAGAAACAAGATGTC 2009
QY 2169 CGGAGTTTGTGAGAAAGGGGGCGCTGTGCCCTGCCAGAGCTGTGCTGATGCCGTG 2228
DB 2010 ATCGGGGTCTGAGAAAGAGAAAGACCGGCTGCCCAAGCTGATCTGTGTCAACGGTCTT 2069
QY 2229 TTCAGGCTCATGAGAGTGTGAGCTTATGAGCTTGGGAGCGGCGCCAGCTTACG 2285
DB 2070 TATACCTCATGACCCGCTGTGGAGCTTACGACCCCAAGTGAAGCGGCCCGCTTACCC 2126

Search completed: March 19, 2005, 14:24:13
Job time : 442.254 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 03:22:22 ; Search time 7723.14 Seconds
(without alignment)

13179.075 Million cell updates/sec

Title: US-10-660-763-1

Perfect score: 2674
Sequence: 1 tccggcgccgcgcacccggcc.....aaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_hlc:*
4: gb_esc3:*
5: gb_esc4:*
6: gb_esc5:*
7: gb_esc6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1806.8	67.6	2461	3	CR624741 full-length
2	1189	44.5	1189	3	CR593957 full-length
3	923.6	34.5	936	5	BX356089 BX356089
4	908.8	34.0	980	5	BX356088 BX356088
5	897.6	33.6	3373	3	AK004587 Mus muscu
6	885.4	33.1	936	1	AL542976 AL542976
7	875.8	32.8	1029	4	BM456755 AGENCOURT
8	863	32.3	863	3	CR622908 full-length
9	855.2	32.0	989	5	BX359009 BX359009
10	836.6	31.3	861	1	AL540776 AL540776
11	821.6	30.7	948	5	BX359010 BX359010
12	805	30.1	868	1	AL540775 AL540775
13	804.4	30.1	939	5	BX381726 BX381726
14	791.6	29.6	913	1	AL569718 AL569718
15	771.6	28.9	948	5	BQ708270 AGENCOURT
16	731	27.3	732	5	BQ015500 UI-H-E11-
17	716.8	26.8	922	5	BUI56855 AGENCOURT
18	710.6	26.6	1114	4	BM543870 AGENCOURT
19	703.2	26.3	847	5	BUI84736 AGENCOURT
20	681.2	25.5	688	5	BQ016080 UI-H-DT1-
21	679	25.4	734	1	AI862686 wj27903.x
22	671.6	25.1	735	5	BQ711139 AGENCOURT
23	669.4	25.0	973	5	BX327963 BX327963
24	652.8	24.4	802	5	BP157303 BP157303

25	640.2	23.9	787	6	CB990282	CB990282 AGENCOURT
26	640	23.9	659	5	BM929822	BM929822 UI-E-E11-
27	629	23.5	629	7	CV030188	CV030188 9223 Full
28	628.8	23.5	1077	5	BM922603	BM922603 AGENCOURT
29	613.8	23.0	882	6	BY704806	BY704806 BY704806
30	613.4	22.9	880	5	BX381725	BX381725 BX381725
31	606	22.7	638	5	BQ807439	BQ807439 NISC JK04
32	603	22.5	629	4	BM684264	BM684264 UI-E-E11-
33	600.8	22.5	652	7	CN298839	CN298839 170004240
34	600	22.4	602	6	CA449888	CA449888 UI-H-E11-
35	599.4	22.4	962	5	BQ679512	BQ679512 AGENCOURT
36	597	22.3	938	5	BQ712558	BQ712558 AGENCOURT
37	593.8	22.2	678	2	AW130256	AW130256 xE47B03.x
38	588	22.0	782	7	CK357411	CK357411 AGENCOURT
39	587.6	22.0	911	5	BUI91031	BUI91031 AGENCOURT
40	581.6	21.8	812	7	CK788677	CK788677 AGENCOURT
41	578.4	21.6	580	5	BP332887	BP332887 BP332887
42	576.4	21.6	674	7	CK943999	CK943999 4068060 B
43	574.8	21.5	912	5	BQ924403	BQ924403 AGENCOURT
44	568	21.2	582	5	BP256574	BP256574 BP256574
45	551.2	20.6	662	2	AW701492	AW701492 uq86h02.y

ALIGNMENTS

RESULT 1	CR624741	2461 bp	mRNA	linear	HTC 21-JUL-2004
LOCUS	CR624741				
DEFINITION	full-length cDNA clone CS0D1069YK21 of Placenta Cot 25-normalized				
ACCESSION	CR624741.1	GI:50505548			
VERSION	CR624741.1	GI:50505548			
KEYWORDS	HTC; CNSLT_cDNA.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Bukacinska; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jesse, J., and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
REMARK	Contact : Feng Liang Email : fliang@lifeitech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paraday Avenue 2 (bases 1 to 2461)				
REFERENCE	Genoscope.				
AUTHORS	Direct Submission				
TITLE	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage :				
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)				
COMMENT	- Web : www.genoscope.cns.fr				
	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.				
FEATURES	Location/Qualifiers				
source	1..2461				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CS0D1069YK21"				
	/issue_type="Placenta Cot 25-normalized"				
	/plasmid="pCMVSPORT_6"				
ORIGIN					
Query Match	67.6%; Score 1806.8; DB 3; Length 2461;				
Best Local Similarity	85.3%; Pred. No. 2.1e-311;				
Matches 2205; Conservative	0; Mismatches 2; Indels 377; Gaps				
	2;				
Qy	41	CCAGAGCAGCTGCCGTCGGAACGACGACTTGGCTTCTTCTGAGCTGACGCC			
Db	1	CCAGAGCAGCTGCCGTCGGAACGACGACTTGGCTTCTTCTGAGCTGACGCC			

QY 101 CCAGGCGCA CGGGGCTCTG CAGCAAA TGCAGAGGCCGAGCTT CTTCTA CTGAGGGCA T 160
Db 61 CCAGGCGCA CGGGGCTCTG CAGCAAA TGCAGAGGCCGAGCTT CTTCTA CTGAGGGCA T 120
QY 161 GAGAAAGTGATGCC CAGCGGGTCAAGATGA CAGAGATATGAGAGCTGCTTCA CCA 220
Db 121 GAGAAAGTGATGCC CAGCGGGTCAAGATGA CAGAGATATGAGAGCTGCTTCA CCA 180
QY 221 CATGTCCCTG CAGAGAGATGGGGGCCAGAGCCCGGGCATCA GGCCTGAGACCCCATCA G 280
Db 181 CATGTCCCTG CAGAGAGATGGGGGCCAGAGCCCGGGCATCA GGCCTGAGACCCCATCA G 240
QY 281 TCAGTCTGGGCTGAGATCA CAGCCAACTGAGGGCTGAGCGGCTTGCTGCGGAGCA 340
Db 241 TCAGTCTGGGCTGAGATCA CAGCCAACTGAGGGCTGAGCGGCTTGCTGCGGAGCA 300
QY 341 CGCAGAGATCTGAACTCAGGGGCCCTGAGCAAGCTGAGCTGCTATCCGGGAACGGCA 400
Db 301 CGCAGAGATCTGAACTCAGGGGCCCTGAGCAAGCTGAGCTGCTATCCGGGAACGGCA 360
QY 401 GCAGCTTCCGAGACCTTACAGCGAGCATGTGGCAGAGCTG CAGAGAGCTTCA CCAAGAC 460
Db 361 GCAGCTTCCGAGACCTTACAGCGAGCATGTGGCAGAGCTG CAGAGAGCTTCA CCAAGAC 420
QY 461 CCACAGCCAGAGACATTTGAGAAAGCTGAAAGCCAGTACCGAGCTTGGCA CCGGAGCA GTGC 520
Db 421 CCACAGCCAGAGACATTTGAGAAAGCTGAAAGCCAGTACCGAGCTTGGCA CCGGAGCA GTGC 480
QY 521 CCAGAGCCAGAGATCCAGAGAGCCAGCAAGCA CAAAGACCGTGA CCAAGGCCAAGGA 580
Db 481 CCAGAGCCAGAGATCCAGAGAGCCAGCAAGCA CAAAGACCGTGA CCAAGGCCAAGGA 540
QY 581 CAAATATGTGGCAGAGCTGTGAAAGCTCTTTGCTACCA CACCGCTATGTGTGGGCGT 640
Db 541 CAAATATGTGGCAGAGCTGTGAAAGCTCTTTGCTACCA CACCGCTATGTGTGGGCGT 600
QY 641 GCGGGCTGCGCAGCTTACACCA CAGCAGCA CCA CACAGCTCTGCTCGCGGCTGCTGCG 700
Db 601 GCGGGCTGCGCAGCTTACACCA CAGCAGCA CCA CACAGCTCTGCTCGCGGCTGCTGCG 660
QY 701 GTCACTGCAAGAGCTTGCACAGAGAGATGTGCTTCATCTGAAAGAGATCTG CAGAGAA T 760
Db 661 GTCACTGCAAGAGCTTGCACAGAGAGATGTGCTTCATCTGAAAGAGATCTG CAGAGAA T 720
QY 761 CCTGAGATTAAGAGCTGTGTGAGAGATGAGGTGTGGCCATTCACCGGAGATGTGCTG 820
Db 721 CCTGAGATTAAGAGCTGTGTGAGAGATGAGGTGTGGCCATTCACCGGAGATGTGCTG 780
QY 821 AGCTGCTGCCGATCCAGCTGAGCTGAGCTGAGTACCAAGGCTTCTGCGACAGATGAGGTC 880
Db 781 AGCTGCTGCCGATCCAGCTGAGCTGAGCTGAGTACCAAGGCTTCTGCGACAGATGAGGTC 840
QY 881 CGCAGCTGAGCTTCCACCTGTGTCA CAGTTCATGAGTCACTGCTTGAGAGGGTGAAC 940
Db 841 CGCAGCTGAGCTTCCACCTGTGTCA CAGTTCATGAGTCACTGCTTGAGAGGGTGAAC 900
QY 941 GCTGAGAGCTGGGAGAGCTCAGCTGAA CAGAGCTGAGCTGAGAGAGGCTGACACAGCT 1000
Db 901 GCTGAGAGCTGGGAGAGCTCAGCTGAA CAGAGCTGAGCTGAGAGAGGCTGACACAGCT 960
QY 1001 GACCTAGTGA CAGATGAGCTGCTGAGCCACCGAGATGTGTTCA GGGCGGAGAGAT 1060
Db 961 GACCTAGTGA CAGATGAGCTGCTGAGCCACCGAGATGTGTTCA GGGCGGAGAGAT 1020
QY 1061 GGTTAAGCAGCTCA CAGAGAGCTCCGGAATGAAGAGAGAACCA CCA CCGCGGAGAG 1120
Db 1021 GGTTAAGCAGCTCA CAGAGAGCTCCGGAATGAAGAGAGAACCA CCA CCGCGGAGAG 1080
QY 1121 GGTGAGAGCTGCTGGCAGAGAGGAGCATGTGCTGCAAGAGCACTG CAGGGGCTG CAGGTAGC 1180
Db 1081 GGTGAGAGCTGCTGGCAGAGAGGAGCATGTGCTGCAAGAGCACTG CAGGGGCTG CAGGTAGC 1140

QY 1181 GCTGTGACCCAGGCCAAGCTG CAGGCC CAGAGAGATGTGCTG CAGACCAAGCTGAGCA 1240
Db 1141 GCTGTGACCCAGGCCAAGCTG CAGGCC CAGAGAGATGTGCTG CAGACCAAGCTGAGCA 1200
QY 1241 CTTGGGCCCCCGAGAGCCCGGCTGTGTCTGCTTGCAGAGATGACCGGCACTTCCAGCTC 1300
Db 1201 CTTGGGCCCCCGAGAGCCCGGCTGTGTCTGCTTGCAGAGATGACCGGCACTTCCAGCTC 1260
QY 1301 GTCTTCGAGCAGAGAGCAGAGAGGAGAGAGACCA CCA CAGCTGAGATCTTTAAGAGCA 1360
Db 1261 GTCTTCGAGCAGAGAGCAGAGAGGAGAGAGACCA CCA CAGCTGAGATCTTTAAGAGCA 1320
QY 1361 CATCTCAGGAATCTTCCGCCCCAAGTCTC ----- 1390
Db 1321 CATCTCAGGAATCTTCCGCCCCAAGTCTC GCTCTCCCTC CACCGCTG CAGCTCATTC CCGA 1380
QY 1391 ----- 1390
Db 1381 GGTGAGAGCCCTG CAGTGA CAGCTGTGTATCCA CCGGGGCATTC CAGAGG CAGAGGT 1440
QY 1391 ----- 1390
Db 1441 GGCTGAGCTGCTGTGCACTTGGGAGCTTCTGTGGTGGAGAGCC CAGAGG CAGAGGA 1500
QY 1391 ----- 1390
Db 1501 GTACGTGCTGTGGTGTGGGATGTGCTGCCCGGCACTT CATCATC CAGGCTCAGCG 1560
QY 1391 ----- GAACTGTATCCGACTG 1406
Db 1561 CCCCTCAGATGAGAGCTGCTGACCCCGGGTCCCTGCCCTG CAGAACTGTATCCGACTG 1620
QY 1407 GAAAGGAGAGGCTTCTTCTGAGCACTTCTTGTCTATGACCACTACTGAGCA CCGCAGAG 1466
Db 1621 GAAAGGAGAGGCTTCTTCTGAGCACTTCTTGTCTATGACCACTACTGAGCA CCGCAGAG 1680
QY 1467 CCCCTCAGCAAGAGATGTGTGTCTGCA CAGGGCTGTGCCCAAGCAAGATGGGTG 1526
Db 1681 CCCCTCAGCAAGAGATGTGTGTCTGCA CAGGGCTGTGCCCAAGCAAGATGGGTG 1740
QY 1527 CTGAACCATGAGAGCTGTGTGTGGGTGAGCATTTGGA CCGGGGAACTTTGGCGGAAGTG 1586
Db 1741 CTGAACCATGAGAGCTGTGTGTGGGTGAGCATTTGGA CCGGGGAACTTTGGCGGAAGTG 1800
QY 1587 TTCAGCGAGCGCTGCGAGCCGACAA CACCTGTGTGGCTGTAAGTCTTGTGAGAGAGC 1646
Db 1801 TTCAGCGAGCGCTGCGAGCCGACAA CACCTGTGTGGCTGTAAGTCTTGTGAGAGAGC 1860
QY 1647 CTCACA CTTGACCTCAGAGCCAA GTTTCTTACAGAA GCGAGATCTTGAAGCATGACG 1706
Db 1861 CTCACA CTTGACCTCAGAGCCAA GTTTCTTACAGAA GCGAGATCTTGAAGCATGACG 1920
QY 1707 CACCCCAATATGTGTGTCTTCA TTGTGTGTGCA CCAAGAGAGCCATCTTACATCGTC 1766
Db 1921 CACCCCAATATGTGTGTCTTCA TTGTGTGTGCA CCAAGAGAGCCATCTTACATCGTC 1980
QY 1767 ATGAGAGCTTGTGAGGGGGGCACTTCTGACCTTCTCGGCA CCGAGGGGGGCCGCTG 1826
Db 1981 ATGAGAGCTTGTGAGGGGGGCACTTCTGACCTTCTCGGCA CCGAGGGGGGCCGCTG 2040
QY 1827 CGGGTGAAGACTGTG CAGATGTGGGAGTGCAGCTGTG CACTGAGATCCTGAG 1886
Db 2041 CGGGTGAAGACTGTG CAGATGTGGGAGTGCAGCTGTG CACTGAGATCCTGAG 2100
QY 1887 AGCAAGTGTGATCCA CCGGAGCTGTGCTGCTCGGA CTGCTGTGTG CAGAGAAAT 1946
Db 2101 AGCAAGTGTGATCCA CCGGAGCTGTGCTGCTCGGA CTGCTGTGTG CAGAGAAAT 2160
QY 1947 GTCTGGAAGATCAGTGA CTTTGGGATGTCCGAGAGAAAGCCGATGGGGCTTATG CAGCC 2006
Db 2161 GTCTGGAAGATCAGTGA CTTTGGGATGTCCGAGAGAAAGCCGATGGGGCTTATG CAGCC 2220
QY 2007 TCAGGGGGCTCAGCAAGTCCCGTGAAGTGA CCGCACTGAGGCCCTTAACTACGGC 2066

Db 2221 TCAGGGGGGCTCAGACAAATCCCCGTAAGTGAACCGCACCTGAGCCCTTAACTAC--- 2277
Qy 2067 CGCTACTCTCCGAAAGCAGACGTGTGAGCTTTGGCATCTTGTCTGGAGAACCTTACAGC 2126
Db 2278 ----- 2277
Qy 2127 CTGGGGGGCTCCCCCTATCCCAACTCAGCAATCAGACACACGGAGATTGTGAGAAAG 2186
Db 2278 ----- 2277
Qy 2187 GGGGGGCGCTGTGCGCTGCGCAGAGCTGTGCTGATGCGGTGTCAGGCTCATGAGACAG 2246
Db 2278 GGGGGGCGCTGTGCGCTGCGCAGAGCTGTGCTGATGCGGTGTCAGGCTCATGAGACAG 2337
Qy 2247 TGCTGGGCTATGAGCTGTGGGACCGGCCAGCTTGACCACTTACAGAGAGCTGACG 2306
Db 2338 TGCTGGGCTATGAGCTGTGGGACCGGCCAGCTTGACCACTTACAGAGAGCTGACG 2397
Qy 2307 AGCATTCGAAAGCGGCATGTGTGAGGCTGGGACCCCTTCTCAAGCTGTGGCTCTGCA 2366
Db 2398 AGCATTCGAAAGCGGCATGTGTGAGGCTGGGACCCCTTCTCAAGCTGTGGCTCTGCA 2457
Qy 2367 GGCC 2370
Db 2458 GGCC 2461

RESULT 2
CR593957 1189 bp mRNA linear HMC 21-JUL-2004
LOCUS full-length cDNA clone CS0D1008Y07 of Placenta Cot 25-normalized
DEFINITION
ACCESSION CR593957
VERSION CR593957.1 GI:50474764
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1189)
L1 W.B., Gruber C., Jesse J., and Polyes D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1189)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source 1. 1189 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1008Y07"
/cissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 44.5%; Score 1189; DB 3; Length 1189;
Best Local Similarity 100.0%; Pred. No. 2.1e-201;
Matches 1189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1326 GGAAGACACCCACGCTGAGATCTTAAAGCACATCTCAGAAATCTTCCGCCCAAG 1385

Db 1 GGAAGACACCCACGCTGAGATCTTAAAGCACATCTCAGAAATCTTCCGCCCAAG 60
Qy 1386 TTCTGGAAGCTGTACCGATCGTGAAGGGAAGGCTTTCTTACGATCTCTTGTCTCATGAC 1445
Db 61 TTCTGGAAGCTGTACCGATCGTGAAGGGAAGGCTTTCTTACGATCTCTTGTCTCATGAC 120
Qy 1446 CACCTTACAGACACCCAGACGCCCTCACCAAGAAAGTGTGTCTCTGCAAGAGGCT 1505
Db 121 CACCTTACAGACACCCAGACGCCCTCACCAAGAAAGTGTGTCTCTGCAAGAGGCT 180
Qy 1506 GTGCCAAGACAAAGTGGTGTCTGAAACATGAGGACCTGTGGTGGAGACAGATTGA 1565
Db 181 GTGCCAAGACAAAGTGGTGTCTGAAACATGAGGACCTGTGGTGGAGACAGATTGA 240
Qy 1566 GGGGGGAACCTTGGGGAAGTGTTCAGCGGACGCTTGAGCCGACACACACCTGTGTGGG 1625
Db 241 GGGGGGAACCTTGGGGAAGTGTTCAGCGGACGCTTGAGCCGACACACACCTGTGTGGG 300
Qy 1626 GTGAAGCTTTGTGAGAGACGCTCCGACCTTACAGGCCAAGTTTCTACAGAAAGCG 1685
Db 301 GTGAAGCTTTGTGAGAGACGCTCCGACCTTACAGGCCAAGTTTCTACAGAAAGCG 360
Qy 1686 AGGATCTGAGAGAGTACAGCCACCCCAATGTCGTCTCATTTGTGTCTGACCCAG 1745
Db 361 AGGATCTGAGAGAGTACAGCCACCCCAATGTCGTCTCATTTGTGTCTGACCCAG 420
Qy 1746 AAGCAGCCCATCTACATCTGTATGAGCTTGTGACAGGGGGGAGCTTCTGACCTTCTC 1805
Db 421 AAGCAGCCCATCTACATCTGTATGAGCTTGTGACAGGGGGGAGCTTCTGACCTTCTC 480
Qy 1806 CGCAGGAGGGGGCCCGCTGGGGGTGAAGACTTGTCTGCAATGTGGGGGATGCAAGCT 1865
Db 481 CGCAGGAGGGGGCCCGCTGGGGGTGAAGACTTGTCTGCAATGTGGGGGATGCAAGCT 540
Qy 1866 GCTGGCATGAGATACCTGAGAGACAAAGTGTGATCCACCGGACCTGTGCTCGGAAAC 1925
Db 541 GCTGGCATGAGATACCTGAGAGACAAAGTGTGATCCACCGGACCTGTGCTCGGAAAC 600
Qy 1926 TGCTGTGTAAGAGAGAAATGTCTTGAAGATCAGTATGATGTGTGAGATGTCGAGAGAA 1985
Db 601 TGCTGTGTAAGAGAGAAATGTCTTGAAGATCAGTATGATGTGTGAGATGTCGAGAGAA 660
Qy 1986 GCCGATGGGATCTATGACAGCTTCAAGGGGGCTTCAAGCAAGTCCCCGTGAAGTGAACCGCA 2045
Db 661 GCCGATGGGATCTATGACAGCTTCAAGGGGGCTTCAAGCAAGTCCCCGTGAAGTGAACCGCA 720
Qy 2046 CCGTAGGCTTAACTAAGAGGCGCTTCTCCTCGAAAGGAGCGTGTGAGACTTGTGATC 2105
Db 721 CCGTAGGCTTAACTAAGAGGCGCTTCTCCTCGAAAGGAGCGTGTGAGACTTGTGATC 780
Qy 2106 TTGCTGTGAGAGACCTTACAGCTGTGGGGGCTCCCTATCCCACTCAGCAATCAGCAG 2165
Db 781 TTGCTGTGAGAGACCTTACAGCTGTGGGGGCTCCCTATCCCACTCAGCAATCAGCAG 840
Qy 2166 ACAAGGAGTTGTGAGAGAGGGGGGCGCTGTGCTGCTGCGCAGAGCTGTGTCTGATGCC 2225
Db 841 ACAAGGAGTTGTGAGAGAGGGGGGCGCTGTGCTGCTGCGCAGAGCTGTGTCTGATGCC 900
Qy 2226 GTGTTAGGCTCATGAGAGAGTGTGGGCTTATGAGCTGTGGGACGGGCGAGCTTACAG 2285
Db 901 GTGTTAGGCTCATGAGAGAGTGTGGGCTTATGAGCTGTGGGACGGGCGAGCTTACAG 960
Qy 2286 ACCATCTACAGAGAGCTGAGAGACATCCGAAAGCGGACATCGTGTGAGCTGTGAGCCCTT 2345
Db 961 ACCATCTACAGAGAGCTGAGAGACATCCGAAAGCGGACATCGTGTGAGCTGTGAGCCCTT 1020
Qy 2346 CTCAGCTGTGTGCTGTGAGAGCTGTGAGAGCTTCTCAAGGGGCTCCAGCTCATATGC 2405
Db 1021 CTCAGCTGTGTGCTGTGAGAGCTGTGAGAGCTTCTCAAGGGGCTCCAGCTCATATGC 1080
Qy 2406 TGACAGCTCTTCAAGTCTGTGAGCTGTGCAACAGATCCACATGCGCGGAGAGATGCA 2465

Db	1081	TGACAGCTCTTACAGACTCCTTGAGACTCCCTGCCACCAAGATTCACACTGCCGGCAGAGTGC	1149
QY	2466	GCGCCGTGTCTCTCTGTGTGTCTCTGCTGCTGCCAAGGCTTCTCTCCG	2514
Db	1141	GCGCCGTGTCTCTCTGTGTGTCTCTGCTGCTGCCAAGGCTTCTCTCCG	1189
RESULT 3			
BX356089			
LOCUS			
DEFINITION	BX356089 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CS0D1008YG07 5-PRIME, mRNA sequence.		
ACCESSION	BX356089		
VERSION	BX356089.2		
KEYWORDS	EST.		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Li,W.B., Gruber,C., Jeesee,J. and Polayes,D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	On May 5, 2003 this sequence version replaced gi:30372027. Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-0ligo(dT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7663.f For more information about this cluster, see http://www.genoscope.cns.fr/cdnatse=CS0D1008AD04QP1tc=7663.f .		
FEATURES			
Source			
	1..936		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CS0D1008YG07"		
	/tissue_type="PLACENTA COT 25-NORMALIZED"		
	/clone_1lb="Homo sapiens PLACENTA COT 25-NORMALIZED"		
	/note="1st strand cDNA was primed with a NotI-0ligo(dT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."		
ORIGIN			
Query Match	34.5%; Score 923.6; DB 5; Length 936;		
Best Local Similarity	99.8%; Pred. No. 3.7e-154;		
Matches 934; Conservative	1; Mismatches 0; Indels 1; Gaps 1;		
QY	1326	GGAGGACACCCACGCTGGAATCTTTAAAGCCACATCTCAGCATTTCCGCCCAAG	1385
Db	1	GGAGGACACCCACGCTGGAATCTTTAAAGCCACATCTCAGCATTTCCGCCCAAG	60
QY	1386	TTCTGCACTGTACCAGCTGGAAGGGAAGGCTTCTTAGCATCTCTTGCTCATGAC	1445
Db	61	TTCTGCACTGTACCAGCTGGAAGGGAAGGCTTCTTAGCATCTCTTGCTCATGAC	120
QY	1446	CACCTACTGAGCACCCACAGCCCCCTCACCAAGAAGATGATTGTCTGACAGGGCT	1505
Db	121	CACCTACTGAGCACCCACAGCCCCCTCACCAAGAAGATGATTGTCTGACAGGGCT	180
QY	1506	GTGCCAAGACAAGTGGGTGTGAACCATGAGACTGTGTTGGTGAACAATTGGA	1565
Db	181	GTGCCAAGACAAGTGGGTGTGAACCATGAGACTGTGTTGGTGAACAATTGGA	240
QY	1566	CGGGGAACCTTTGGGGAAGTGTTCAGGGAGCCTTCGAGCCGACACAACCTTGATGCG	1623
Db	241	CGGGGAACCTTTGGGGAAGTGTTCAGGGAGCCTTCGAGCCGACACAACCTTGATGCG	300

Oy	1626	GTGAAGCTTGTGTGAGAGAGCGCTCCACCTTGACTCTTAAGGCCAAGTTTCTACAGAAAGCG	16693
Db	301	GTGAAGCTTGTGTGAGAGAGCGCTCCACCTTGACTCTTAAGGCCAAGTTTCTACAGAAAGCG	360
Oy	1686	AGGATCTTGAAGACAGTACAGCCACCACCAATCGTGCCTCATTTGGTGTGTGCACCCAG	17451
Db	361	AGGATCTTGAAGACAGTACAGCCACCACCAATCATGTGTGTCTCATTGGTGTGTGCACCCAG	420
Oy	1746	AAGCAGCCCATCTACATGTCATGAGAGCTTGTGCAGGGGGGAGCGACTTCTGACCTTCTCT	18051
Db	421	AAGCAGCCCATCTACATGTCATGAGAGCTTGTGCAGGGGGGAGCGACTTCTGACCTTCTCT	480
Oy	1806	CGCAGCGGAGGGGGGGCGCGCTGCGGGGTGAAGAATCTGTGTGCAGATAGTGGGGGATGACAGCT	18651
Db	481	CGCAGCGGAGGGGGGGCGCGCTGCGGGGTGAAGAATCTGTGTGCAGATAGTGGGGGATGACAGCT	540
Oy	1866	GCTGGCATGAGATACCTGAGAGAGCAAGTGTGCATCCACCGGAGACTGGCTGTCCGAGAC	19251
Db	541	GCTGGCATGAGATACCTGAGAGAGCAAGTGTGCATCCACCGGAGACTGGCTGTCCGAGAC	600
Oy	1926	TGCTGTGTGAACAGAAAGAAATGTCTGAAATCATGTACATTTGGGATGTCCCGAGAGAA	19851
Db	601	TGCTGTGTGAACAGAAAGAAATGTCTGAAATCATGTACATTTGGGATGTCCCGAGAGAA	660
Oy	1986	GCCGATGGGGTCTATGAGAGCTCAGAGGGGGCTCAGACAGTCCCGTGAAGTGAACGCA	20451
Db	661	GCCGATGGGGTCTATGAGAGCTCAGAGGGGGCTCAGACAGTCCCGTGAAGTGAACGCA	720
Oy	2046	CCTGAGGCCCTTAACTACAGGCCGCTACTCTCTCGAAAGCAAGTGTGAAGCTTTGGCATC	21051
Db	721	CCTGAGGCCCTTAACTACAGGCCGCTACTCTCTCGAAAGCAAGTGTGAAGCTTTGGCATC	780
Oy	2106	TTGCTCTGGGAGACCTTACAGCCTTGGGGGCTCCCCCTAATCCCAACTCAGCAATCAGCA	21641
Db	781	TTGCTCTGGGAGACCTTACAGCCTTGGGGGCTCCCCCTAATCCCAACTCAGCAATCAGCA	840
Oy	2165	GACACGGGAGTTTGTGAGAAAGGGGGGCGCTCTGCGCTGCGCCAGAGCTGTCTCTGATGC	22241
Db	841	GACACGGGAGTTTGTGAGAAAGGGGGGCGCTCTGCGCTGCGCCAGAGCTGTCTCTGATGC	900
Oy	2225	CGTGTTCAGGCTCATGAGCAGTGTCTGGGCTTATGA	2260
Db	901	CGTGTTCAGGCTCATGAGCAGTGTCTGGGCTTATGA	936
RESULT 4			
LOCUS	EX356088/c	980 bp	RNA linear EST 08-APR-2004
DEFINITION	EX356088 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA		
ACCESSION	EX356088		
VERSION	EX356088.2		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 980)		
JOURNAL	L4,W.B., Gruber,C., Jesssee,J. and Polayes,D.		
COMMENT	Full-length CDNA libraries and normalization		
	Unpublished (2001)		
	On May 5, 2003 this sequence version replaced gi:30369992.		
	Contact: Genoscope		
	Genoscope - Centre National de Sequenage		
	2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr		
	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers		
	end enriched, double-strand cDNA was digested with Not I and cloned		
	into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library		
	was normalized. Library was constructed by Life Technologies, a		
	division of Invitrogen. This sequence belongs to sequence cluster		
	7663.F		

For more information about this cluster, see
<http://www.genoscope.cne.fr/cdna?cs=CSODI008AD04NP1ec=7663.f>.

FEATURES
 Location/Qualifiers

1..980
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODI008Y007"
 /clone_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime and enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 34.0%; Score 908.8; DB 5; Length 980;

Best Local Similarity 95.2%; Pred. No. 1.6e-151;

Matches 926; Conservative 24; Mismatches 21; Indels 2; Gaps 2;

1533 CATGAGAGCTGGTGTGGTGAGCAGATTGACGCGGGAATTGGCGAAGTGTTCAGC 1592
 980 CATGAGAGCTGGTGTG-TGGGTGAGCAGATTGACGCGGGAATTGGCGAAGTGTTCAGC 922
 1593 GGAAGCGCTGCGAGCGCAACACCCGTGGTGGCGGTGAAGTCTTGAGAGAGCGTCCCA 1652
 921 GGAAGCGCTGCGAGCGCAACACCCGTGGTGGCGGTGAAGTCTTGAGAGAGCGTCCCA 862
 1653 CCTGACCTCA-GGCGAAGTTTCTACAGAGAGAGAGATCTCGAAGCAGTACAGCACC 1711
 861 CCTGACCTCAAGGCGCAAGTTTCTACAGAGAGAGAGATCTCGAAGCAGTACAGCACC 802
 1712 CAACATGTCGCTCTCATTTGGTGTCTGCAACCCAGAGAGAGAGATCTCATGTCATGGA 1771
 801 CAACATGTCGCTCTCATTTGGTGTCTGCAACCCAGAGAGAGAGATCTCATGTCATGGA 742
 1772 GCTGTGCAAGGGGGGAGCTTCTGACCTTCCCTCCGCAAGAGGGGGCCCGCTGGGGGT 1831
 741 GCTGTGCAAGGGGGGAGCTTCTGACCTTCCCTCCGCAAGAGGGGGCCCGCTGGGGGT 682
 1832 GAAGACTCTGCTGCAAGTGTGGGGATGCGAGTGTGGCATGAGTACCTGAGAGACA 1891
 681 GAAGACTCTGCTGCAAGTGTGGGGATGCGAGTGTGGCATGAGTACCTGAGAGACA 622
 1892 GTGTGCTGCTGCAAGTGTGGGGATGCGAGTGTGGCATGAGTACCTGAGAGACA 1951
 621 GTGTGCTGCTGCAAGTGTGGGGATGCGAGTGTGGCATGAGTACCTGAGAGACA 562
 1952 GAAGACTCTGCTGCAAGTGTGGGGATGCGAGTGTGGCATGAGTACCTGAGAGACA 2011
 561 GAAGACTCTGCTGCAAGTGTGGGGATGCGAGTGTGGCATGAGTACCTGAGAGACA 502
 2012 GGGGCTCAGACAGTCCCGTGAAGTGAACCGGACCTGAGAGCCCTTAACCTAGCGGCGCTA 2071
 501 GGGGCTCAGACAGTCCCGTGAAGTGAACCGGACCTGAGAGCCCTTAACCTAGCGGCGCTA 442
 2072 CTCCTCCGAAGGAGCGTGTGAGCTTTGGCATCTTGTCTTGAGAGACCTTCAGCTGGG 2131
 441 CTCCTCCGAAGGAGCGTGTGAGCTTTGGCATCTTGTCTTGAGAGACCTTCAGCTGGG 382
 2132 GGGCTTCCCTTATCCCACTCAGCAATCAGCAACCGGGAATTTGGAGAGAGGGGG 2191
 381 GGGCTTCCCTTATCCCACTCAGCAATCAGCAACCGGGAATTTGGAGAGAGGGGG 322
 2192 CCGTCTGCGCTGCGCAAGAGTGTGTCTGATGCGGTGTGAGGCTCATGAGAGAGTGTG 2251
 321 CCGTCTGCGCTGCGCAAGAGTGTGTCTGATGCGGTGTGAGGCTCATGAGAGAGTGTG 262
 2252 GGGCTTATGAGCGCTGAGCAGCGCCAGCTTACAGACATCTACAGAGAGTGTGAGAGCAT 2311
 261 GGGCTTATGAGCGCTGAGCAGCGCCAGCTTACAGACATCTACAGAGAGTGTGAGAGCAT 202
 2312 CCGAAGGCGGATCGGTGAGGCTGGAGACCCCTTCTCAAGCTGTGTGAGGCTCTGAGAGCCT 2371

DB 201 CCGAAGGCGGATCGGTGAGGCTGGAGACCCCTTTTAAGCTGTGTGCTCTGAGAGCCT 142

QY 2372 AGGTGAGCTCTCTCAGCGGCTCAGCTCATATGCTGACAGCTCTTTCACAGTCTGAGCTC 2431

DB 141 AGGTGAGCTCTCTCAGCGGCTCAGCTCATATGCTGACAGCTCTTTCACAGTCTGAGAGC 82

QY 2432 CTGCGACCGAGTCCACATCTGCGCGCAGAGATGACGCGCGGTCTCTGTGTCTGCTG 2491

DB 81 CTGCGACCGAGTCCACATCTGCGCGCAGAGATGACGCGCGGTCTCTGTGTGTGCGNNN 22

QY 2492 TGTGTGAGGAGCT 2504

DB 21 NTATGCGAGGTT 9

RESULT 5
 AK004587 3373 bp mRNA linear HTC 03-APR-2004
 LOCUS
 DEFINITION Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200003015 product:similar to PROTO-ONCOGENE TYROSINE-PROTEIN KINASE FES/FPS (BC 2.7.1.112) (C-FES) [Mus musculus], full insert sequence.

ACCESSION AK004587.1 GI:12835858
 VERSION AK004587
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20493374
 PUBMED 11042159

REFERENCE
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Katsunai, T., Taahito, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multiplexed capillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 MEDLINE 11076861

REFERENCE
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 MEDLINE 11076861

REFERENCE
 AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Aikawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,

Db 1429 CAGAGCCCAAGCCCTGCTCTTTCAGATGACCGCCATTCACCTCTCCACGAGCA 1488

Qy 1313 GAGCGAGAGGGGGAGAGACACCCAGCTGAGATCTTTAAGAGCCACATTCAGAAAT 1372

Db 1489 GGAGCGAGAGGGGGAGAGACACCCAGCTGAGATCTTTAAGAGCCACATTCCTCGAAT 1548

Qy 1373 CTTCCGCCCAAGTTTCGACACTGTACCGACTGGA 1408

Db 1549 CTTCCGCCCAAGTTTCGACACTGTACCGACTGGA 1584

RESULT 6
AL542976 936 bp mRNA linear EST 24-MAR-2004
LOCUS AL542976 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013YP02
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION AL542976
VERSION AL542976.3 GI:45718543
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 936)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:30548660.
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seq@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 7663.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0DE013DH01QP1c=7663.f.
FEATURES
Source 1..936
Location/Qualifiers
1..936
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE013YP02"
/issue_type="PLACENTA"
/clone_id="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 33.1%; Score 885.4; DB 1; Length 936;
Best Local Similarity 98.8%; Pred. No. 2.4e-147;
Matches 900; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

Qy 29 TCCGAGGCGCTCCGAGGACAGCTGCCCTGCGGACACACATATAGGCTTCTCTTCA 88

Db 14 TCCCGGAGATGTCGAGAGCAGCTGCCCTGCGGACACACATATAGGCTTCTCTTCA 73

Qy 89 GCTGTGACAGCCCGCAGGCGCACCGGGTCTGTGACGAATATGACGAGCCGAGCTTGTCT 148

Db 74 GCTGTGACAGCCCGCAGGCGCACCGGGTCTGTGACGAATATGACGAGCCGAGCTTGTCT 133

Qy 149 ACTGAGGCGCATGAGAAAGTGAATGCGCCAGCGGTCAAGAGTGACAGGAGATATGACAG 208

Db 134 ACTGAGGCGCATGAGAAAGTGAATGCGCCAGCGGTCAAGAGTGACAGGAGATATGACAG 193

Qy 209 ACTGCTTACACATCTCCCTGACGACAGATGGGGGCCAGAGCCGGCCATCAGCCCTGA 268

Db 194 ACTGCTTACACATCTCCCTGACGACAGATGGGGGCCAGAGCCGGCCATCAGCCCTGA 253

Qy 269 CAGCCCATCAGTCACTGCTGAGATCAACAGCCAACTGAGGGCTTGAAGCCGTT 328

Db 254 CAGCCCATCAGTCACTGCTGAGATCAACAGCCAACTGAGGGCTTGAAGCCGTT 313

Qy 329 GCTGGGCGAGCAGCAGAGATCTGAATCAAGGGCCCTGAGCAAGCTGAGCTGTCT 388

Db 314 GCTGGGCGAGCAGCAGAGATCTGAATCAAGGGCCCTGAGCAAGCTGAGCTGTCT 373

Qy 389 CCGGAGAGCGCAGCAGCTTCGCAAGACTCAAGCAGCAGTGGCAGAGCTCAGAGGA 448

Db 374 CCGGAGAGCGCAGCAGCTTCGCAAGACTCAAGCAGCAGTGGCAGAGCTCAGAGGA 433

Qy 449 GCTCAACCAAGCCCAAGCAGCAGATTTGAAGAGCTGAAGAGCCAGTACCGAGCTTGGC 508

Db 434 GCTCAACCAAGCCCAAGCAGCAGATTTGAAGAGCTGAAGAGCCAGTACCGAGCTTGGC 493

Qy 509 ACGGACAGTGGCCCAAGCCCAAGCTCAAGAGAGGCGCAGCAAGCAAGAGCCGTGA 568

Db 494 ACGGACAGTGGCCCAAGCCCAAGCTCAAGAGAGGCGCAGCAAGCAAGAGCCGTGA 553

Qy 569 CAAGGCAAGAGCAAGATATGTCGAGAGCTTGAAGAGCTTGTCTCAACCAAGCGCTA 628

Db 554 CAAGGCAAGAGCAAGATATGTCGAGAGCTTGAAGAGCTTGTCTCAACCAAGCGCTA 612

Qy 629 TGTGCTGGGCGTGGCGGCTGCGCAGCTACACCAAGCAGCAGCAGCTCTGTGCG 688

Db 613 TGTGCTGGGCGTGGCGGCTGCGCAGCTACACCAAGCAGCAGCAGCTCTGTGCG 672

Qy 689 CCGCTGCTGCGGCTGCTGCTGAGAGACTTGAAGAGCTTGTCTCAACCAAGCGCTA 748

Db 673 CCGCTGCTGCGGCTGCTGCTGAGAGACTTGAAGAGCTTGTCTCAACCAAGCGCTA 732

Qy 749 CCGTCAAGAAATCTGAGATTTAGCAGCCTGTGAGAGTGTGTGGCCATTCAAGC 808

Db 733 CCGTCAAGAAATCTGAGATTTAGCAGCCTGTGAGAGTGTGTGGCCATTCAAGC 792

Qy 809 GAGATAGCTGACAGCTGTGCGGATCCAGCTTGAAGCTGAGTCAAGAGCTTCTGCG 868

Db 793 GAGATAGCTGACAGCTGTGCGGATCCAGCTTGAAGCTGAGTCAAGAGCTTCTGCG 852

Qy 869 ACAGTATGGTCCGACCTGAGAGTCCACCTGTGTCAAGTCACTGCTTGA 928

Db 853 ACAGTATGGTCCGACCTGAGAGTCCACCTGTGTCAAGTCACTGCTTGA 912

Qy 929 GGAGGGTGAAC 939

Db 913 GGAGGGTGAAC 923

RESULT 7
BM456755 1029 bp mRNA linear EST 05-FEB-2002
LOCUS BM456755
DEFINITION AGENCOURT_6403957 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5583652
5', mRNA sequence.
ACCESSION BM456755
VERSION BM456755.1 GI:18505795
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1029)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-rc@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM12347 row: b column: 05
 High quality sequence stop: 704.
 Location/Qualifiers

FEATURES

source

1..1029
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5583652"
 /issue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH MGC 92"
 /notes="Organ: testis; Vector: pCMV-Sport6; Site: 1: NotI;
 Site: 2: SalI; Cloned unidirectionally; oligo-dt primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH MGC library."

ORIGIN

Query Match 32.8%; Score 875.8; DB 4; Length 1029;
 Best Local Similarity 98.0%; Pred. No. 1.2e-145;
 Matches 939; Conservative 0; Mismatches 13; Indels 6; Gaps 5;

2 CCGGGGTCGGCAACCGGGCTGAGTCCGTCGAGCCGTCGAGACAGCTCCGCTGGG 61
 11 CCGGGGTCGGCAACCGGGCTGAGTCCGTCGAGCCGTCGAGACAGCTCCGCTGGG 70
 62 GAACAGACTATGGGCTTCTCTCTGAGTGTGACAGCCCGGAGGCGGAGGCTCTGCA 121
 71 GAACAGACTATGGGCTTCTCTCTGAGTGTGACAGCCCGGAGGCGGAGGCTCTGCA 130
 122 GCAATGTCAGAGAGCGGAGCTTCTCTCTGAGTGTGACAGCCCGGAGGCGGAGGCT 181
 131 GCAATGTCAGAGAGCGGAGCTTCTCTCTGAGTGTGACAGCCCGGAGGCGGAGGCT 190
 182 GGTCAAGATGACAGGAGATGACAGAGCTTCAACATGTCCTCGGAGGAGCTGAG 241
 191 GGTCAAGATGACAGGAGATGACAGAGCTTCAACATGTCCTCGGAGGAGCTGAG 250
 242 GGGCCAGAGCCGGGCGCATCAGCCCTGACAGCCCATCAGTCTCTGGGCTGAGATCAC 301
 251 GGGCCAGAGCCGGGCGCATCAGCCCTGACAGCCCATCAGTCTCTGGGCTGAGATCAC 310
 302 CAGCCAACTGAGAGGCGCTTGTGCTGCGGAGCAGCAGAGAGATCTGAATCTCAG 361
 311 CAGCCAACTGAGAGGCGCTTGTGCTGCGGAGCAGCAGAGAGATCTGAATCTCAG 370
 362 GGGCCAGAGCTGAGGCTGCTCATCGGGAGAGCGGAGGAGCTTGGAGGCTTACAG 421
 371 GGGCCAGAGCTGAGGCTGCTCATCGGGAGAGCGGAGGAGCTTGGAGGCTTACAG 430
 422 CGAGAGTGGCAGCAGCTGACAGAGAGCTCACAAGACCCAGCAGGACATTGAGAA 481
 431 CGAGAGTGGCAGCAGCTGACAGAGAGCTCACAAGACCCAGCAGGACATTGAGAA 490
 482 GCTGAAGAGCAGTACCGAGCTTGTGAGCGGAGAGCTGCGCAAGCCAGGAGTACCA 541
 491 GCTGAAGAGCAGTACCGAGCTTGTGAGCGGAGAGCTGCGCAAGCCAGGAGTACCA 550
 542 GAGAGCAGCAAGAGACCGGTGACAAAGGCGCAAGGACAAAGTATGTGGGAGCTGTG 601
 551 GAGAGCAGCAAGAGACCGGTGACAAAGGCGCAAGGACAAAGTATGTGGGAGCTGTG 610
 602 GAAGCTCTTGTGCTCACCACACCGCTATGTGCTGGGCTGCGGAGCTGCGGAGTACCA 661
 611 GAAGCTCTTGTGCTCACCACACCGCTATGTGCTGGGCTGCGGAGCTGCGGAGTACCA 670
 662 CCAGACACACACAGCTCTGTGCTGCGGAGCTGCTGCGGAGCTGCGGAGCTGCGGAG 721
 671 CCAGACACACACAGCTCTGTGCTGCGGAGCTGCTGCGGAGCTGCGGAGCTGCGGAG 730

QY 722 GGAGATGGCTTCATCTGTAAGAGATCTGAGGAATACCTGAGATTAGAGCTGTG 781
 DB 731 GGAGATGGCTTCATCTGTAAGAGATCTGAGGAATACCTGAGATTAGAGCTGTG 790
 QY 782 GCAGATGAGTGTGTGGCATTACCGGG-AGATGGCTGACAGCTGTGCGGATCAGC 840
 DB 791 GCAGATGAGTGTGTGGCATTACCGGGAGATGAGCTGACAGCTGTGCGGATCAGC 850
 QY 841 CTGAGGCTGAGTACCAAGGCTTCTGCGACATTAATGGTCCGCACTGAC-CTGCCACCC 899
 DB 851 CTGAGGCTGAGTACCAAGGCTTCTGCGACATTAATGGTCCGCACTGACGCTCAACCC 909
 QY 900 TGTGTACG-TTCGATGAGTCACTGCTTGA--GAGAGGTGAACCGCTGAGGCTGGG 954
 DB 910 TGTGTACGTTTCGATGAGTCACTGCTTGAAGAGGGGTGAACCGCTGAGGCTGGG 967

RESULT 8
 CR622908 863 bp mRNA linear HTC 21-JUL-2004
 LOCUS
 DEFINITION
 full-length cDNA clone CSDB004YN09 of Placenta of Homo sapiens
 (human).
 accession CR622908
 version CR622908.1 GI:50503715
 keywords HTC; CNSLT cDNA.
 source Homo sapiens (human)
 organism Homo sapiens

REFERENCE
 AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
 Faraday Avenue
 2 (bases 1 to 863)
 REFERENCE
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 Location/Qualifiers

FEATURES
 source 1..863
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSDB004YN09"
 /issue_type="Placenta"
 /plasmid="pCMVSPORT_6"

ORIGIN

Query Match 32.3%; Score 863; DB 3; Length 863;
 Best Local Similarity 100.0%; Pred. No. 2.3e-143;
 Matches 863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1686 AAGATCTGAACACATGACCAACCCCAATGCGCTCATTTGATGTCACCCAG 1745
 DB 1 AAGATCTGAACACATGACCAACCCCAATGCGCTCATTTGATGTCACCCAG 60
 QY 1746 AAGCAGCCCATATACATGCTATGAGCTTGTGCAAGGGGGGAGCTTCTGACTTCTC 1805
 DB 61 AAGCAGCCCATATACATGCTATGAGCTTGTGCAAGGGGGGAGCTTCTGACTTCTC 120
 QY 1806 CGCAGAGAGGGGGCCGCTGCGGAGTGAAGACTCTGCTGACAGATGGTGGGAGTACGCT 1865
 DB 121 CGCAGAGAGGGGGCCGCTGCGGAGTGAAGACTCTGCTGACAGATGGTGGGAGTACGCT 180

QY 1866 GCTGGATGAGTAACTTGAAGAGCAAGTGTGATCCACCGGACCTGTGTGCTCGGAAC 1925
 DB 181 GCTGGATGAGTAACTTGAAGAGCAAGTGTGATCCACCGGACCTGTGTGCTCGGAAC 240
 QY 1926 TGGCTGCTGACAGAGAAAGATGTCTTGAAGATCACTTGGAGTGTCCGAGAGAA 1985
 DB 241 TGGCTGCTGACAGAGAAAGATGTCTTGAAGATCACTTGGAGTGTCCGAGAGAA 300
 QY 1986 GCCGATGGGAGTCTATGCAAGCTTCAAGGGGCTTCAAGCAAGTCCCGTGAAGTGAACCGA 2045
 DB 301 GCCGATGGGAGTCTATGCAAGCTTCAAGGGGCTTCAAGCAAGTCCCGTGAAGTGAACCGA 360
 QY 2046 CCTGAGGCCCCCTTAACATACGCGCCCTACTCTCTCGAAAGGACGCTGTGAGCTTTGGCATC 2105
 DB 361 CCTGAGGCCCCCTTAACATACGCGCCCTACTCTCTCGAAAGGACGCTGTGAGCTTTGGCATC 420
 QY 2106 TTGCTCTGGAGAGCTTCAAGCTTCAAGGGGCTTCCCTTATCCAACTCCAGCAATCAAGCAG 2165
 DB 421 TTGCTCTGGAGAGCTTCAAGCTTCAAGGGGCTTCCCTTATCCAACTCCAGCAATCAAGCAG 480
 QY 2166 ACAAGGAGTGTGAGAGAGAGGGGCGCTGTGCTGCTGCGCAGAGCTGTGTCTGATGCC 2225
 DB 481 ACAAGGAGTGTGAGAGAGAGGGGCGCTGTGCTGCTGCGCAGAGCTGTGTCTGATGCC 540
 QY 2226 GTGTTAGAGCTATGAGAGAGTGTGAGGCTTATGAGCTTGGGAGAGGCGCCAGCTTCAAGC 2285
 DB 541 GTGTTAGAGCTATGAGAGAGTGTGAGGCTTATGAGCTTGGGAGAGGCGCCAGCTTCAAGC 600
 QY 2286 ACCATCTACAGAGAGCTGAGAGCATCCGAAAGCGGATGATGAGAGCTGGAGAGCCCTT 2345
 DB 601 ACCATCTACAGAGAGCTGAGAGCATCCGAAAGCGGATGATGAGAGCTGGAGAGCCCTT 660
 QY 2346 CTCAGCTGTGAGGCTCTGTGAGGCTTATGAGCTTCTCAAGGCTTCAAGCTCATATGC 2405
 DB 661 CTCAGCTGTGAGGCTCTGTGAGGCTTATGAGCTTCTCAAGGCTTCAAGCTCATATGC 720
 QY 2406 TGAAGCTCTTCAAGCTTCTGTGAGCTTCTGTGAGAGATCAAGATCAAGTCCGAGAGATGCA 2465
 DB 721 TGAAGCTCTTCAAGCTTCTGTGAGCTTCTGTGAGAGATCAAGATCAAGTCCGAGAGATGCA 780
 QY 2466 GAGCGCTGTCTCTGTGAGCTTCTGTGAGCTTCTGTGAGAGATCAAGATCAAGTCCGAGAGATGCA 2525
 DB 781 GAGCGCTGTCTCTGTGAGCTTCTGTGAGCTTCTGTGAGAGATCAAGATCAAGTCCGAGAGATGCA 840
 QY 2526 TAAACCACTTGTGAGCACTGAA 2548
 DB 841 TAAACCACTTGTGAGCACTGAA 863

RESULT 9
 BX359009/c 989 bp mRNA linear EST 08-APR-2004
 LOCUS BX359009 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CS0D1051YO20 3-PRIME, mRNA sequence.

ACCESSION BX359009.2 GI:46305682
 VERSION BX359009.2
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 989)
 Li, W.B., Gruber, C., Jessee, D., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 5, 2003 this sequence version replaced gi:30376296.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
 Email: sequef@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 7663.f

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?cs=CS0D1051BH1ONP1c=7663.f>.

FEATURES

Location/Qualifiers

1..989

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D1051YO20"

/issue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 32.0%; Score 855.2; DB 5; Length 989;

Best Local Similarity 95.8%; Pred. No. 5,6e-142;

Matches 944; Conservative 14; Mismatches 18; Indels 9; Gaps 8;

QY 1510 CCAAGGACAGTGGGCTGCAACCATGAGAGACCTGTGTGGGTGAGCAGATTGACGGG 1569
 DB 989 CCAAGGACAGTGGGCTGCAACCATGAGAGACCTGTGTGGGTGAGCAGATTGACGGG 994
 QY 1570 GGAACCTTTGGGCGAAGGTTCAAGCGAGCGCTGCGAGCCGACCAACACC-CTGTGTGCGGTG 1628
 DB 933 GGAACCTTTGGGCGAAGGTTCAAGCGAGCGCTGCGAGCCGACCAACACC-CTGTGTGCGGTG 874
 QY 1629 AAGTCTTGTGAGAGAGCGTCCCACTGACCTTCAAGGCGCAAGTTTCAAGAGAGAGG 1688
 DB 873 AAGTCTTGTGAGAGAGCGTCCCACTGACCTTCAAGGCGCAAGTTTCAAGAGAGAGG 814
 QY 1689 ATCTGGAACAGATGACGACCAACCAATGCGGCTCATTTGTGTCTGACCCGAGAG 1748
 DB 813 ATCTGGAACAGATGACGACCAACCAATGCGGCTCATTTGTGTCTGACCCGAGAG 754
 QY 1749 CAGCCC-ATCTGATGATGATGAGAGCTTGTGAGAGGAGGCGACCTTCTGACCTTCTCCG 1807
 DB 753 CAGCCC-ATCTGATGATGATGAGAGCTTGTGAGAGGAGGCGACCTTCTGACCTTCTCCG 694
 QY 1808 CAGGAGGGGGCGCGCTGCGGGTGAAGACTTGTCTGACAGATGTGGGGATGACCTGC 1867
 DB 693 CAGGAGGGGGCGCG-CTGCGGGTGAAGACTTGTCTGACAGATGTGGGGATGACCTGC 635
 QY 1868 TGGCATGAGATCACTGAGAGAGCAAGTGTGATCCACC-AGGACCTGCTGCTCGGAACCT 1926
 DB 634 TGGCATGAGATCACTGAGAGAGCAAGTGTGATCCACC-AGGACCTGCTGCTCGGAACCT 575
 QY 1927 GCTGTGACAGAGAAAGATGTCTTGAAGATCACTTGGAGATGTCCCGAGAGAGA-A 1985
 DB 574 GCTGTGACAGAGAAAGATGTCTTGAAGATCACTTGGAGATGTCCCGAGAGAGAGA 515
 QY 1986 GCCGATGGGAGTCTATGCAAGCTTCAAGGGGCTTCAAGCAAGTCCCGTGAAGTGAACCGCA 2045
 DB 514 GCCGATGGGAGTCTATGCAAGCTTCAAGGGGCTTCAAGCAAGTCCCGTGAAGTGAACCGCA 455
 QY 2046 CCTGAGGCCCCCTTAACATACGCGCCCTACTCTCGAAAGGACGCTGTGAGCTTTGGCATC 2105
 DB 454 CCTGAGGCCCCCTTAACATACGCGCCCTACTCTCGAAAGGACGCTGTGAGCTTTGGCATC 395
 QY 2106 TTGCTCTGGAGAGCTTCAAGCTTCAAGGGGCTTCCCTTATCCAACTCCAGCAATCAAGCAG 2165
 DB 394 TTGCTCTGGAGAGCTTCAAGCTTCAAGGGGCTTCCCTTATCCAACTCCAGCAATCAAGCAG 335
 QY 2166 ACAAGGAGTGTGAGAGAGAGGGGCGCTGTGCTGCTGCGCAGAGCTGTGTCTGATGCC 2225
 DB 334 ACAAGGAGTGTGAGAGAGAGGGGCGCTGTGCTGCTGCGCAGAGCTGTGTCTGATGCC 275
 QY 2226 GTGTTAGAGCTATGAGAGAGTGTGAGGCTTATGAGCTTGGGAGGCGCCAGCTTCAAGC 2285

Db 274 GTGTCAGGCTCATGAGCAGTGTCTGGCCATGAGCCTGGGACGCGCBACGCTTCACG 215
Qy 2286 ACCATTACCAAGAGAGCTGCGAGAGATCCGAAAGCGGATCGGTGAGGCTGGGACCCCTT 2345
Db 214 ACCATTACCAAGAGAGCTGCGAGAGATCCGAAAGCGGATCGGTGAGGCTGGGACCCCTT 155
Qy 2346 CTCAGCTGTGAGCTCTGCGAGGCTTAGTGAGAGCTCTCAGGCGGCTCCAGCTCATATGC 2405
Db 154 CTCAGCTGTGAGCTCTGCGAGGCTTAGTGAGAGCTCTCAGGCGGCTCCAGCTCATATGC 95
Qy 2406 TGAACAGCTTTCACAGCTCTGAGCTCTGCGACACAGATCCACATGCGCGGAGATGCA 2465
Db 94 TGAACAGCTTTCACAGCTCTGAGCTCTGCGACACAGATCCACATGCGCGGAGATGCA 35
Qy 2466 GCGCGCTGTCTCTCTGTGTCTCTG 2490
Db 34 GCGCGCTGTCTCTCTGTGTCTCTG 10

RESULT 10
AL540776 861 bp mRNA linear EST 24-MAR-2004
LOCUS AL540776 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE004YN09
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION AL540776
VERSION AL540776.3 GI:45716373
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 861)
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:30544306.
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 7663.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?b=CS0DE004CG05QPI&c=7663.f.
FEATURES
source
1..861
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE004YN09"
/issue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 31.3%; Score 836.6; DB 1; Length 861;
Best Local Similarity 99.4%; Pred. No. 1,le-138;
Matches 858; Conservative 2; Mismatches 1; Indels 2; Gaps 2;
Qy 1686 AGGATCTGAGAGAGTACAGCAGCCCAACATGCTGCTCTCACTTGTCTCCACACCG 1745
Db 1 AGGATCTGAGAGAGTACAGCAGCCCAACATGCTGCTCTCACTTGTCTCCACACCG 60
Qy 1746 AACGACCCATCTACATCTCATGAGCTTGTGCAAGGGGCGCACTTCTGACCTTCCTC 1805

Db 61 AAGCAGCCCATTTATCATCGTATGAGCTTGTGCAAGGGGCGAGACTTCTGACCTTCCTC 120
Qy 1806 CGCAGAGAGAGGGGCGCGCTGGGGGTGAAGACTCTGCGCAATGCTGGGGGATGAGAGT 1865
Db 121 CGCAGAGAGAGGGGCGCGCTGGGGGTGAAGACTCTGCGCAATGCTGGGGGATGAGAGT 180
Qy 1866 GCTGAGATGAGATACCTGAGAGCAAGTGTGATTCACACCGGGAACCTGTGCTCGGAAC 1925
Db 181 GCTGAGATGAGATACCTGAGAGCAAGTGTGATTCACACCGGGAACCTGTGCTCGGAAC 240
Qy 1926 TGCCTGTGACAGAGAGAGATGTCTGAAGATGATGACTTTGGATGTCCGAGAGAA 1985
Db 241 TGCCTGTGACAGAGAGAGATGTCTGAAGATGATGACTTTGGATGTCCGAGAGAA 300
Qy 1986 GCGGATGGGCTTATGAGAGCTTCAGGGGCTTCAGACAAAGTCCCGTGAAGTGAACCGCA 2045
Db 301 GCGGATGGGCTTATGAGAGCTTCAGGGGCTTCAGACAAAGTCCCGTGAAGTGAACCGCA 360
Qy 2046 CCTGAGAGCCCTTAATACAGGCGGCTACTCTCCGAAGGAGAGTGTGAGACTTTGGCATC 2105
Db 361 CCTGAGAGCCCTTAATACAGGCGGCTACTCTCCGAAGGAGAGTGTGAGACTTTGGCATC 420
Qy 2106 TTGCTCTGGAGACCTTCAGGCTGAGGCGCTTCCTCCATTCACAACTTCAGCAATCAGCA 2165
Db 421 TTGCTCTGGAGACCTTCAGGCTGAGGCGGCTTCCTCCATTCACAACTTCAGCAATCAGCA 480
Qy 2166 AACAGGAGTTTGTGAGAGAGGGGGCGGCTGCGCTGCGCAAGAGTGTCTGATGCGC 2225
Db 481 AACAGGAGTTTGTGAGAGAGGGGGCGGCTGCGCTGCGCAAGAGTGTCTGATGCGC 540
Qy 2226 GTGTTACAGCTCATGAGAGAGTGTGAGGCTTATGAGCTGGGAGAGCGCCAGCTTCACG 2285
Db 541 GTGTTACAGCTCATGAGAGAGTGTGAGGCTTATGAGCTGGGAGAGCGCCAGCTTCACG 600
Qy 2286 ACCATCTACAGAGAGCTGAGAGATCCGAAGCGGATCGGTGAGCTGGGACCCCTT 2345
Db 601 ACCATCTACAGAGAGCTGAGAGATCCGAAGCGGATCGGTGAGCTGGGACCCCTT 659
Qy 2346 CTCAGCTGTGAGCTCTGCGAGGCTTAGTGAGAGCTCTCAGGCGCTCCAGCTCATATGC 2405
Db 660 CTCAGCTGTGAGCTCTGCGAGGCTTAGTGAGAGCTCTCAGGCGCTCCAGCTCATATGC 719
Qy 2406 TGAACAGCTTTCACAGCTCTGAGCTCTGCGACACAGATCCACATGCGCGGAGATGCA 2465
Db 720 TGAACAGCTTTCACAGCTCTGAGCTCTGCGACACAGATCCACATGCGCGGAGATGCA 779
Qy 2466 GCGCGCTGTCTCTGTGTCTCTGCTGCTGCGCAGGCGCTTCTTCGCGGCAAGAACAA 2525
Db 780 GCGCGCTGTCTCTGTGTCTCTGCTGCTGCGCAGGCGCTTCTTCGCGGCAAGAACAA 839
Qy 2526 TAAACCACTTGTGCCCACTGAA 2548
Db 840 T-AAACCACTTGTGCCCACTGAA 861

RESULT 11
BX359010 968 bp mRNA linear EST 08-APR-2004
LOCUS BX359010 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1051YO20 5-PRIME, mRNA sequence.
ACCESSION BX359010
VERSION BX359010.2 GI:46307488
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 968)
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30378289.

[illegible]

RESULT 14	
AL569718/c	
LOCUS	913 bp mRNA linear EST 05-APR-2004
DEFINITION	AL569718 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013YP02
ACCESSION	AL569718
VERSION	AL569718.3 GI:46235981
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE	1 (bases 1 to 913)
AUTHORS	Li, W.B., Gruber, C., Jessup, J., and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	On Feb 16, 2001 this sequence version replaced gi:30606435.

Genoscope - Centre National de Séquençage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqrel@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-Oligo (dT) primer. Five primers
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 7663.f
For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?b=CS05D011DHO1NP1&c=7663.f>.

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS00E013YP02"
/tissue_type="PLACENTA"
/clone_id="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

	Query Match	Score	DB	Length
	29.6%	791.6	DB 1	913
	Beat Local Similarity	94.5%	Pred. No. 1,2e-110	
	Matches	850	Conservative 17	Mismatches 27
			Indels 5	Gaps 5
Qy	1593	GGAGCGCTCGACGCCGACACACCTCGTGGCGGTGAAGTC-TTGTGAGAGAGAGCTTCC	1651	
Db	913	GGAGCGCTCGACGCCGACACACCTCGTGGCGGTGAAGTC-TTGTGAGAGAGAGCTTCC	854	
Yy	1652	ACCTGACCTCAAGGCCAAGTTT-CTACGAGAGGAGAGAGATCTGAGCAGTACAGCCACC	1710	

Db	853	ACCGACACTCAAGGCGCAAGTTTCTTACAGGAAGGAGAGATCTGAAGCAGGACAGCCACC	794
OY	1711	CCAAACATCGTGCAGTCAATGGTGTGTGTCGCAACCCAGAAAGCAGCCCATCTACATCGTCATGG	1770
Db	793	CCAAACATGTCGCTCTCATTTGGTGTCTGCAACCCAGAAAGCAGCCCATCTACATCGTCATGG	734
OY	1771	AGCTTGTGACAGAGGGGGGCGACTTCTCTGACCTTCTCCGACCGAGGGGGGCCGCGCTCGGGG	1830
Db	733	AGCTTGTGACAGAGGGGGGCGACTTCTCTGACCTTCTCCGACCGAGGGGGGCCGCGCTCGGGG	675
OY	1831	TGAAGACTCTGCTGCAATGTGTGGGGGATGCACTGCTGGGATGGAAGTACTTGGAGAGCA	1890
Db	674	TGAAGACTCTGCTGCAATGTGTGGGGGATGCACTGCTGGGATGGAAGTACTTGGAGAGCA	615
OY	1891	AGTCTGTCATCCACCGGGACCTGGCTGTGCGAATCGCCTGAGCAGAGAAATGTC	1949
Db	614	AGTCTGTCATCCACCGGGACCTGGCTGTGCGAATCTGATCTGTAACARBAATAATTTCC	555
OY	1950	CTGAAGATCAAGTGACTTTGGAGATGCCGAGAGAGAGCCGATGGGGGTCTATGACGCTCA	2009
Db	554	CTGAAGATCAAGTGACTTTGGAGATTTCCCBAGAGAGAGCCGATGGGGGTCTATGACGCTCA	495
OY	2010	GGGGGCGCTCAGACAAAGTCCCGCTGAAGTGCACCGACCTGAGGCCCTTAACTACGGCCGC	2068
Db	494	GGGGGCGCTCAGACAAATTTCCCGTGAAGTGCACCCACCTGAGGCCCTTAACTACGGCCGC	435
OY	2070	TACTCTCCCGAAACGAGTGTGAGGCTTTGGACATTTGCTCTGGAGAGACCTTACAGCGTG	2128
Db	434	TACTCTCCCGAAACCTAGTGTGAGGCTTTGGACATTTGCTCTGGAGAGACCTTACAGCGTG	375
OY	2130	GGGGGCGTCCCTTATCCCACTTCAGACATTCAGACACACGGGAGTTTGTGGAGAGGGG	2188
Db	374	GGGGGCGTCCCTTATCCCACTTCAGACATTCAGACATTCAGACACGGGAGTTTGTGGAGAGGGG	315
OY	2190	GGCGGTCTGCCCTTCCCAAGCTGTGTCTGTATGCCGTTTCAAGGCTCATGGAAGCATGTC	2248
Db	314	GGCGGTCTCCCTTCCCAAGCTGTGTCTGTATGCCGTTTCAAGGCTCATGGAAGCATGTC	255
OY	2250	TGGGCGTATGAGCGCTGGGCGAGCGGCGCCAGCTTACAGACCATCTACAGAGACCTGACAGC	2309
Db	254	TGGGCGTATGAGCTTGGGCGAGCGGCGCCAGCTTACAGACCATCTACAGAGACCTGACAGC	195
OY	2310	ATCCGAAAGCGGACATCGGTAGAGCTGGGACCCCTTCTCAAGCTGTGTGGCTCTGACAGC	2365
Db	194	ATCCGAAAGCGGACATCGGTAGAGCTGGGACCCCAATAAATCTGTGTGCAATCTGACAGC	135
OY	2370	CTAAGGTGAGCTCTTCAAGCGGCTTCAGCTCATATGTGTACAGCTTTTCAAGTCTTGAC	2428
Db	134	CTAAGGTGAGCTCTTCAAGCGGCTTCAGCTCATATGTGTACAGCTTTTCAAGTCTTGAC	75
OY	2430	TCTGTGCCACCAAGATTCACATGCGGGGAGGATGACAGCGCGTGTCTGTGTGTCTCC	2488
Db	74	YCCGTGCACCAATATCCAACTGCGCGGAGGATTCAGACGCGT-TCTCATCTGTNCCC	17

[illegible]

Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Mark Watson
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLCM2516 row: o column: 07
 High quality sequence stop: 610.

FEATURES

Source

Location/Qualifiers
 1. 948
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /mol_type="mRNA"
 /clone="IMAGE:6301278"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH_MGC_113"
 /note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

ORIGIN

Query Match 28.9%; Score 771.6; DB 5; Length 948;
 Best Local Similarity 95.3%; Pred. No. 4.3e-127;
 Matches 818; Conservative 0; Mismatches 34; Indels 6; Gaps 2;

```

QY 294 GAGATACACGACCAAACTGAGGGCTGAGCGCTTGCTGGCGACGACGAGAGATCTG 353
DB 1 GAGATACACGACCAAACTGAGGGCTGAGCGCTTGCTGGCGACGACGAGAGATCTG 60
QY 354 AACTCAGGGCCCTTGAGCAAGCTGAGCTGCTCATCCGGGACGGCAGCAGCTTCGCAAG 413
DB 61 AACTCAGGGCCCTTGAGCAAGCTGAGCTGCTCATCCGGGACGGCAGCAGCTTCGCAAG 120
QY 414 ACCTACAGGAGAGTGGGACGAGCTGCGACGAGAGAGTCAACAAACCCACAGCCGAGAC 473
DB 121 ACCTACAGGAGAGTGGGACGAGCTGCGACGAGAGAGTCAACAAACCCACAGCCGAGAC 180
QY 474 ATTGAGAAGGTGAAGGCGCATACCGAGCTTGGSCACGGGACAGTGCACCAAGCCAAAGCGG 533
DB 181 ATTGAGAAGGTGAAGGCGCATACCGAGCTTGGSCACGGGACAGTGCACCAAGCCAAAGCGG 240
QY 534 AAGTACCAAGAGGCGCAAGAAAGCAAGGACCTGACAAAGGCCAAAGACAATATGTGCGG 593
DB 241 AAGTACCAAGAGGCGCAAGAAAGCAAGGACCTGACAAAGGCCAAAGACAATATGTGCGG 300
QY 594 AGCTGTGGAAGCTCTTTGCTCACCAACCGCTATGTGTGGCGTGGGGCTGGCGCAG 653
DB 301 AGCTGTGGAAGCTCTTTGCTCACCAACCGCTATGTGTGGCGTGGGGCTGGCGCAG 360
QY 654 CTACACCAAGAGGACCAACGAGCTCTGCTGCCGCGCTGCTGGCGTCACTGACGAGAC 713
DB 361 CTACACCAAGAGGACCAACGAGCTCTGCTGCCGCGCTGCTGGCGTCACTGACGAGAC 420
QY 714 CTGCAAGAGAGATGCTTGATCTCTGAAGAGATCTGCAAGAAATACCTGAGATTAGC 773
DB 421 CTGCAAGAGAGATGCTTGATCTCTGAAGAGATCTGCAAGAAATACCTGAGATTAGC 480
QY 774 AGCTGTGCAAGATGAGTGTGGCCATTCAACGGGAGATGGCTGAGCTGCTGCCCGC 833
DB 481 AGCTGTGCAAGATGAGTGTGGCCATTCAACGGGAGATGGCTGAGCTGCTGCCCGC 540
QY 834 ATCCAGCTGAGGCTGATGACCAAGGCTTCTGCGACAGATAGGTCGCCACCTGACGTC 893
DB 541 ATCCAGCTGAGGCTGATGACCAAGGCTTCTGCGACAGATAGGTCGCCACCTGACGTC 600
QY 894 CCACCTGTGTACGTTGATGATCACTGCTTGAGGAGGAGTGAACCGCTGAGCCTGGG 953
  
```

```

DB 601 CCACCTGTGTACGTTGATGATCACTGCTTGAGAGGAGTGAACCGCTGAGCCTGGG 660
QY 954 GAGCTCCAGCTGAAAGAGCTGACTGAGAGACGTCGACACAGCTGACTCAGTGACA 1013
DB 661 GAGCTCCAGCTGAAAGAGCTGACTGAGAGACGTCGACACAGCTGACTCAGTGACA 720
QY 1014 GATGAGCTGCTGTGGCCACCGAGATGCTTCAAGCGGC--AGGATGCTTACCGCAGC 1071
DB 721 GATGAGCTGCTGTGGCCACCGAGATGCTTCAAGCGGC--AGGATGCTTACCGCAGC 780
QY 1072 TCCAAACAGAGGCTCCGAAATGAAGCA---GAACCCACCCCGGAGCGGCTGCAG 1127
DB 781 TCCAAACAGAGGCTCCGAAATGAAGAGAGAGAAACCCACCCCGGAGCGGCTGCAG 840
QY 1128 CTGCTGGGCAAGAGGCAA 1145
DB 841 CAGCTGCTTGGGCGCAA 858
  
```

Search completed: March 19, 2005, 14:10:40
 Job time : 7733.14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 01:20:31 ; Search time 1275.71 Seconds
(without alignments)
12408.301 Million cell updates/sec

Title: US-10-660-763-1

Perfect score: 2674
Sequence: 1 tccggggcgcgcacccgggccc.....aaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4330206 seqs, 295970667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16dec04:*
1: geneeqn19808:*
2: geneeqn19908:*
3: geneeqn20008:*
4: geneeqn2001as:*
5: geneeqn2001bs:*
6: geneeqn2002as:*
7: geneeqn2002bs:*
8: geneeqn2003as:*
9: geneeqn2003bs:*
10: geneeqn2003cs:*
11: geneeqn2003ds:*
12: geneeqn2004as:*
13: geneeqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2674	100.0	2674 6	ABA94500 Human pro
2	2674	100.0	2674 8	ABV75413 Human kin
3	2674	100.0	2674 12	ADM28580 Human pro
4	2490.2	93.1	2599 12	ADK71888 Human kin
5	2326.4	87.0	2889 6	AA895001 Human DNA
6	1953	73.0	2623 13	ADR39816 Human kin
7	1552.2	58.0	2680 12	ADH43089 CRAM prot
8	856	32.0	856 11	AAAT70060 fag/fps p
9	375	14.0	2818 11	AD171052 Gene enco
10	374	14.0	527 13	ADQ54298 Novel can
11	362.4	13.6	12263 6	ABK84514 Human CDN
12	362.4	13.6	15297 6	ABA94501 Human pro
13	362.4	13.6	15297 8	ABV75414 Human kin
14	362.4	13.6	15297 12	ADM28582 Human pro
15	328.8	12.3	2584 11	AD171054 Gene enco
16	326.8	12.2	449 9	ACH14291 Human adu
17	296.2	11.1	361 6	AA817052 Human pro
18	296.2	11.1	361 10	AA61604 Human pro
19	296.2	11.1	361 10	ABX14976 Human pro
20	271.6	10.2	1779 8	ACC57665 Drosophi

21	271.6	10.2	3198 4	ABL18403	Ab118403 Drosophi
22	271.6	10.2	3369 4	ABL14457	Ab114457 Drosophi
23	201.2	7.3	3286 12	ADK71885	Adk71885 Human kin
24	201.2	7.5	3875 6	ABZ35414	Abz35414 Human gen
25	199.6	7.5	5616 10	AD608824	Ad608824 Novel DNA
26	199.2	7.4	2778 12	ADP28218	Adp28218 Human sec
27	198	7.4	2955 6	ABV78117	Abv78117 Human eph
28	198	7.4	2955 6	ABZ35693	Abz35693 Human eph
29	198	7.4	2955 6	ABX09936	Abx09936 Human eph
30	198	7.4	2955 6	ABL91658	Ab191658 Human pol
31	198	7.4	3370 6	ABL67952	Ab167952 Ovary can
32	198	7.4	3370 8	ABZ34861	Abz34861 Coding se
33	198	7.4	3370 10	ADD18447	Add18447 Human pro
34	198	7.4	3370 10	ADP38440	Adp38440 Human pro
35	198	7.4	3370 12	ADU75094	Ad75094 Marker ge
36	198	7.4	3370 13	ADQ80244	Adq80244 Eph1 CDNA
37	185.4	6.9	3042 6	ABV78118	Abv78118 Human eph
38	185.4	6.9	3042 6	ABZ35694	Abz35694 Human eph
39	185.4	6.9	3042 6	ABX09937	Abx09937 Human eph
40	185.4	6.9	3042 6	ABP91559	Ab191559 Human pol
41	185.4	6.9	3346 13	ACN42099	Acn42099 Human dia
42	185.4	6.9	3319 13	ACN37278	Acn37278 Tumour-as
43	185.4	6.9	3921 6	ABV94231	Abv94231 Breast ca
44	185.4	6.9	3921 8	ABZ34865	Abz34865 Coding se
45	185.4	6.9	3921 10	ADB37439	Adb37439 Human can

ALIGNMENTS

RESULT 1	ABR94500	ABR94500 standard; CDNA; 2674 BP.
ID	ABR94500	standard; CDNA; 2674 BP.
XX	ABR94500;	
AC	ABR94500;	
XX	09-APR-2002	(first entry)
DT	09-APR-2002	(first entry)
XX	Human proto-oncogene tyrosine kinase encoding cDNA.	
DE	Human proto-oncogene tyrosine kinase encoding cDNA.	
XX	Proto-oncogene tyrosine kinase; potK; tumour; cytostatic; anti-leukemic;	
KW	gene therapy; protein therapy; vaccine; enzymatic-inhibition; human;	
KW	anti-kinase; sg.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FT	5'UTR	1..71
FT	5'UTR	72..2330
FT	CDS	/*tag= b
FT	CDS	/product= "proto-oncogene tyrosine kinase"
FT	3'UTR	2331..2674
FT	3'UTR	/*tag= c
XX	US6340584-B1.	
XX	22-JAN-2002.	
XX	27-MAR-2001; 2001US-00817180.	
XX	27-MAR-2001; 2001US-00817180.	
XX	27-MAR-2001; 2001US-00817180.	
XX	(PEKE) PE CORP NY.	
XX	Gan W, Ye J, Di Francesco V, Beasley EM;	
XX	WPI, 2002-138497/18.	
XX	P-PSDB; ABB07354.	
XX	Nucleic acids encoding a proto-oncogene tyrosine kinase, useful for the	
XX	prevention, diagnosis and treatment of e.g. leukemia and lung tumors.	

PS Claim 1; Fig 1A-B; 49pp; English.

XX The invention provides isolated nucleic acid sequences encoding a proto-oncogene tyrosine kinase (pork). The pork polynucleotides and protein may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate pork expression, such as lung and kidney tumors, leukemia and stomach adenocarcinoma. pork may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of pork by expressing CC inactive proteins or to supplement the patient's own production of pork. CC The encoded pork may be used as an antigen in the production of CC antibodies against pork and in assays to identify modulators of pork CC expression and activity. The anti-pork antibodies and antagonists may be CC used to down regulate expression and activity and as diagnostic agents CC for detecting the presence of pork in samples. The present sequence CC represents a cDNA encoding the human pork

SQ Sequence 2674 BP; 662 A; 767 C; 810 G; 435 T; 0 U; 0 Other;

Query Match 100.0%; Score 2674; DB 6; Length 2674;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 2674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCGGGGTCCGACCGGGCTGAGTCCGAGCCGTCGAGAGCACTGCCCCGTCG 60
DB 1 TCCGGGGTCCGACCGGGCTGAGTCCGAGCCGTCGAGAGCACTGCCCCGTCG 60
QY 61 GGAAACGACTATGAGCTTCTCTTCTGAGCTGTGACGCCGCCGAGGCTCTCTG 120
DB 61 GGAAACGACTATGAGCTTCTCTTCTGAGCTGTGACGCCGCCGAGGCTCTCTG 120
QY 121 AGCAATGAGAGAGCGAGCTTCTCTCTGAGCTGTGAGAGGATGAGATGAGGAGC 180
DB 121 AGCAATGAGAGAGCGAGCTTCTCTCTGAGCTGTGAGAGGATGAGATGAGGAGC 180
QY 181 GGATCAAGAGTGCAGAGAGTATGAGAGCTGCTTCAACCATGTCTCTGAGAGAGT 240
DB 181 GGATCAAGAGTGCAGAGAGTATGAGAGCTGCTTCAACCATGTCTCTGAGAGAGT 240
QY 241 GGGGCGAGAGCGGGGCTGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGT 300
DB 241 GGGGCGAGAGCGGGGCTGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGT 300
QY 301 CCAGCCAACTGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 360
DB 301 CCAGCCAACTGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 360
QY 361 GGGCCCTGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 420
DB 361 GGGCCCTGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 420
QY 421 GCGAGT 480
DB 421 GCGAGT 480
QY 481 AGCTGAAGT 540
DB 481 AGCTGAAGT 540
QY 541 AGGAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 600
DB 541 AGGAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 600
QY 601 GGAAGCTCTTGTCTCAACACACCGCTATGTGCTGGGCTGTGGGCTGTGGGCT 660
DB 601 GGAAGCTCTTGTCTCAACACACCGCTATGTGCTGGGCTGTGGGCTGTGGGCT 660
QY 661 ACCAGACCAACACACAGCTCTGTGCTGGGCTGTGGGCTGTGGGCTGTGGGCT 720
DB 661 ACCAGACCAACACACAGCTCTGTGCTGGGCTGTGGGCTGTGGGCTGTGGGCT 720
QY 721 AGGAGATGAGCTTGTCTGTGAAGAGATCTGTGAGAGATCTGTGAGATCTGTGAG 780
DB 721 AGGAGATGAGCTTGTCTGTGAAGAGATCTGTGAGAGATCTGTGAGATCTGTGAG 780

DB 721 AGGAGATGAGCTTGTCTGTGAAGAGATCTGTGAGAGATCTGTGAGATCTGTGAG 780
QY 781 TGCAGAGATGAGTGTGCTGATTCACCGGAGATGCTGCTGCTGCTGCTGCTGCTG 840
DB 781 TGCAGAGATGAGTGTGCTGATTCACCGGAGATGCTGCTGCTGCTGCTGCTGCTG 840
QY 841 CTGAGAGATGAGTGTGCTGATTCACCGGAGATGCTGCTGCTGCTGCTGCTGCTG 900
DB 841 CTGAGAGATGAGTGTGCTGATTCACCGGAGATGCTGCTGCTGCTGCTGCTGCTG 900
QY 901 GTGTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
DB 901 GTGTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
QY 961 AGCTGAAGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1020
DB 961 AGCTGAAGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1020
QY 1021 TGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB 1021 TGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1081 AGCTCGAGATGAG 1140
DB 1081 AGCTCGAGATGAG 1140
QY 1141 GGAAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
DB 1141 GGAAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
QY 1201 TGCAG 1260
DB 1201 TGCAG 1260
QY 1261 GGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB 1261 GGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
QY 1321 AGGAG 1380
DB 1321 AGGAG 1380
QY 1381 CCAAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
DB 1381 CCAAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
QY 1441 TGCAG 1500
DB 1441 TGCAG 1500
QY 1501 GGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
DB 1501 GGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
QY 1561 TGCAG 1620
DB 1561 TGCAG 1620
QY 1621 TGCAG 1680
DB 1621 TGCAG 1680
QY 1681 AGGAG 1740
DB 1681 AGGAG 1740
QY 1741 CCAG 1800
DB 1741 CCAG 1800
QY 1801 TCTTCGAG 1860
DB 1801 TCTTCGAG 1860

[illegible]

RESULT 2	
ABV75413	
ID	ABV75413 standard; cDNA; 2674 BP.
XX	
AC	ABV75413;
XX	
DT	24-JAN-2003 (first entry)
XX	
DE	Human kinase encoding cDNA.
XX	
KW	Human, kinase; proto-oncogene tyrosine kinase; antiarteriosclerotic;
KW	cytostatic; antiinflammatory; antipsoriatic; gene therapy;
KW	protein kinase; drug screening assay; tissue typing; chromosome 15;
KW	pharmacogenomic analysis; inflammation; cancer; leukaemia; lung tumours;
KW	kidney tumour; stomach adenocarcinoma; arteriosclerosis; psoriasis; gene

XX	ss.
OS	Homo sapiens.
XX	
FT	Key
FT	Location/Qualifiers
FT	5'UTR
FT	/tag= a
FT	CDS
FT	/tag= b
FT	/product= "protein kinase"
FT	3'UTR
FT	/tag= c
PX	
PD	WO200277191-A2.
PD	03-OCT-2002.
PF	27-MAR-2002; 2002MO-USO09325.
PR	27-MAR-2001; 2001US-00817180.
PR	06-DEC-2001; 2001US-00003295.
PA	(PEKE) PE CORP NY.
PI	Gan W, Ye J, Di Francesco V, Beasley EM;
DR	WPI; 2003-029927/02.
XX	P-PSDb; ABB99046.
PT	New human kinase protein, useful for treating or diagnosing disorders associated with an absence of, inappropriate, or unwanted expression of the protein, e.g. inflammation or cancer, in drug screening assays and pharmacogenomics.
PS	Claim 4 (b); Fig 1; 75pp; English.
CC	The invention relates to a newly isolated peptide sequence of a human kinase that is related to the proto-oncogene tyrosine kinase subfamily.
CC	The activity of the kinase of the invention may be described as,
CC	cystostatic, antiatherosclerotic, antiinflammatory and antipsoriatic.
CC	Peptides of the invention are useful in assays to determine the biological activity of the protein, in drug screening assays, tissue typing and pharmacogenomic analysis. They are also useful in treating or diagnosing disorders characterised by an absence of, inappropriate, or unwanted expression of the protein, such as inflammation, cancer (e.g. leukaemia, lung tumours, kidney tumours or stomach adenocarcinoma), arteriosclerosis, and psoriasis. Nucleic acid molecules of the invention are useful as probes, primers and chemical intermediates in biological assays. The peptide and nucleic acid sequences are useful as models for the development of human therapeutic targets, aid in the identification of therapeutic proteins and serve as targets for the development of human therapeutic agents that modulate kinase protein activity in cells and tissues that express the protein. The protein of the invention may also be useful in gene therapy. The gene encoding the protein of the invention has been localised to human chromosome 15. The current sequence represents the human kinase encoding cDNA
SQ	Sequence 2674 BP; 662 A; 767 C; 810 G; 435 T; 0 U; 0 Other; Query Match Best Local Similarity 100.0%; Score 2674; DB 8; Length 2674; Matches 2674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DG	1 TC CGGGGGCGCACCGGGCTGTAGTCGGTTCGAGGCCGTCCACAGAGCAGCTGCCTGCG DB 1 TC CGGGGCTCCCAACCGGGCTGTAGTCGGTTCGAGGCCGTCCACAGAGCAGCTGCCTGCG 60 61 GGAA CAGACTATGGGCTTTCTTTCTGTAGCTGTGCAGGCCCCACAGGCCCACGGGGTCTGCG 120 db 61 GGAA CAGACTATGGGCTTTCTTTCTGTAGCTGTGCAGGCCCCACAGGCCCACGGGGTCTGCG 120 GY 121 AGCAAATGCAGAGGCCGAGCTTGCTGTCTA CTGTGAAGGCGCATAGA AAATGTGATGCCCCAACG 180

D	b	121	AGCAAAATGCAGAGAGCCGACGCTTGCTCATCTGAGAGGGCATATAGAAAGTGGATGGCCAGC	180
Q	y	181	GGGTCAAGAGTGCAGAGGAGTATGCAGAGACTGTTCACACATGTCCCTGCAGGACAGTG	240
D	b	181	GGGTCAAGAGTGCAGAGGAGTATGCAGAGACTGTTCACACATGTCCCTGCAGGACAGTG	240
Q	y	241	GGGGCCAGAGCCCGGGCCCTCAGCCCTTGACACCCCTTAAATGTCATCTCTGGGCTTGAATCA	300
D	b	241	GGGGCCAGAGCCCGGGCCCTCAGCCCTTGACACCCCTTAAATGTCATCTCTGGGCTTGAATCA	300
Q	y	301	CCAGCCAAACTGAGGGGCTTGAGCGCTTGCGGGCAGACGACAGAGGATTTGAATCTCAG	360
D	b	301	CCAGCCAAACTGAGGGGCTTGAGCGCTTGCGGGCAGACGACAGAGGATTTGAATCTCAG	360
Q	y	361	GGCCCTTGAGCAAGCTGAGCCTGTGCATCCGGGAAACGGCAGACGCTTGCAGAACCTTACA	420
D	b	361	GGCCCTTGAGCAAGCTGAGCCTGTGCATCCGGGAAACGGCAGACGCTTGCAGAACCTTACA	420
Q	y	421	GGCAGCAGTGGCAGCAGCTTGACAGAGAGCTCACCAAGACCCACAGCCAGGACATTGAGA	480
D	b	421	GGCAGCAGTGGCAGCAGCTTGACAGAGAGCTCACCAAGACCCACAGCCAGGACATTGAGA	480
Q	y	481	AGCTGAAGAGCGAGTACCGAGCTCTGGCAGGGACAGTGGCCCAAGGCCAAGGCAAGTACC	540
D	b	481	AGCTGAAGAGCGAGTACCGAGCTCTGGCAGGGACAGTGGCCCAAGGCCAAGGCAAGTACC	540
Q	y	541	AGGAGGCCAGCAAAAGACAAAGACCGGTGACAAAGCCCAAGACAAATATGTGGCAGCCTGT	600
D	b	541	AGGAGGCCAGCAAAAGACAAAGACCGGTGACAAAGCCCAAGACAAATATGTGGCAGCCTGT	600
Q	y	601	GGAAGCTCTTTGCTCACAACAACCGCTATGTGTCTTGGGCTGGGGCTTGGCAGCTAACCC	660
D	b	601	GGAAGCTCTTTGCTCACAACAACCGCTATGTGTCTTGGGCTGGGGCTTGGCAGCTAACCC	660
Q	y	661	ACCAAGCACACACACAGAGCTCCGTCGCGCCCGGCTGTGCGGCTCATCTGAGGACCTGGACG	720
D	b	661	ACCAAGCACACACACAGAGCTCCGTCGCGCCCGGCTGTGCGGCTCATCTGAGGACCTGGACG	720
Q	y	721	AGGAGATGGCTTGTCATCTTGAAGAGATCTTGCAGGAATACTTGGAGATTAGACGCTGG	780
D	b	721	AGGAGATGGCTTGTCATCTTGAAGAGATCTTGCAGGAATACTTGGAGATTAGACGCTGG	780
Q	y	781	TGCAGGATGAGAGTGTGTGGCCATTACCCGGGAGATGGCTGCAGCTGTGCCGATCCAGC	840
D	b	781	TGCAGGATGAGAGTGTGTGGCCATTACCCGGGAGATGGCTGCAGCTGTGCCGATCCAGC	840
Q	y	841	CTGAGGCTGAGATCAACAAAGGCTTCCGCGACAGTATGGGTCCGACCTTGACGCTCCACCCCT	900
D	b	841	CTGAGGCTGAGATCAACAAAGGCTTCCGCGACAGTATGGGTCCGACCTTGACGCTCCACCCCT	900
Q	y	901	GTGTCACTTGCATGATGATCACTGTCTTGAAGAGGTGAACCGCTGGAGCCTGGGGAGCTCC	960
D	b	901	GTGTCACTTGCATGATGATCACTGTCTTGAAGAGGTGAACCGCTGGAGCCTGGGGAGCTCC	960
Q	y	961	AGCTGAACGAGCTGACTGTGGAGAGCGTGCACACGCTGACCTCAATGACAAATGAGC	1020
D	b	961	AGCTGAACGAGCTGACTGTGGAGAGCGTGCACACGCTGACCTCAATGACAAATGAGC	1020
Q	y	1021	TGGCTGTGGCCACCGAGATGGTGTTCAGGCGGCGAGAAATGTTATCGAGCTGCACACGG	1080
D	b	1021	TGGCTGTGGCCACCGAGATGGTGTTCAGGCGGCGAGAAATGTTATCGAGCTGCACACGG	1080
Q	y	1081	AGCTCCGGAATGAAGAGAGAACACCCACCCCGGGAGCGGGTGCAGCTGTCTGGGCAAGA	1140
D	b	1081	AGCTCCGGAATGAAGAGAGAACACCCACCCCGGGAGCGGGTGCAGCTGTCTGGGCAAGA	1140
Q	y	1141	GGCAAGTCTCTCAAGAACACTGCACGGGGCTGCAGATGTGCGCTGTGCACGAGCCAAAGC	1200
D	b	1141	GGCAAGTCTCTCAAGAACACTGCACGGGGCTGCAGATGTGCGCTGTGCACGAGCCAAAGC	1200
Q	y	1201	TGCAGGCCACAGAGAGTTGTCTGCAGACCAAGCTTGAAGACACTTGGGCCCCGGCCAGCCCC	1260
D	b	1201	TGCAGGCCACAGAGAGTTGTCTGCAGACCAAGCTTGAAGACACTTGGGCCCCGGCCAGCCCC	1260

QY	1261	CGCTGTGCTGCTCTCTGACAGATGACCGCCACATCCAGTGTCTCTGGACAGAGAGAG	1320
Db	1261	CGCTGTGCTGCTCTCTGACAGATGACCGCCACATCCAGTGTCTCTGGACAGAGAGAG	1320
QY	1321	AGGGGGAGAGACACCCACGCTGGAGATCCTTAAAGACCAATCTCAGAAATCTCCGC	1380
Db	1321	AGGGGGAGAGACACCCACGCTGGAGATCCTTAAAGACCAATCTCAGAAATCTCCGC	1380
QY	1381	CCAAGTCTTGAACCTGTACCGACTGGAAGGGGAAGGCTTTCTTAGATTCCTTTGTCTCA	1440
Db	1381	CCAAGTCTTGAACCTGTACCGACTGGAAGGGGAAGGCTTTCTTAGATTCCTTTGTCTCA	1440
QY	1441	TCGACCACTACTGAGACCCAGACGCCCTCAACCAAGAGATGATGTCTCTGACA	1500
Db	1441	TCGACCACTACTGAGACCCAGACGCCCTCAACCAAGAGATGATGTCTCTGACA	1500
QY	1501	GGGCTGTGCCCAAGAGACAAGTGGGTCTGAACCATGAGACCTGTGTTGGTGTGACAGA	1560
Db	1501	GGGCTGTGCCCAAGAGACAAGTGGGTCTGAACCATGAGACCTGTGTTGGTGTGACAGA	1560
QY	1561	TTTGAACGGGGGAACCTTTGGGGAAGTGTTCAGCGGAGCGCTGAGAGCCGACACACCTGG	1620
Db	1561	TTTGAACGGGGGAACCTTTGGGGAAGTGTTCAGCGGAGCGCTGAGAGCCGACACACCTGG	1620
QY	1621	TGGCGGTGAAGTCTTGTTCGAGAGACGCTCCACCTGACCTCAAGGCCCAAGTTTCTACAG	1680
Db	1621	TGGCGGTGAAGTCTTGTTCGAGAGACGCTCCACCTGACCTCAAGGCCCAAGTTTCTACAG	1680
QY	1681	AAGGAGAGATCTTGAGACAGTACAGCCACCCCAACATCTGTGCTCTCATTTGTGTCTGCA	1740
Db	1681	AAGGAGAGATCTTGAGACAGTACAGCCACCCCAACATCTGTGCTCTCATTTGTGTCTGCA	1740
QY	1741	CCCAAGACAGACCCTATCATCTGTCAATGAGAGCTTGTGACGGGGGGGCACTTCTCTGACT	1800
Db	1741	CCCAAGACAGACCCTATCATCTGTCAATGAGAGCTTGTGACGGGGGGGCACTTCTCTGACT	1800
QY	1801	TCTCTCCGACGAGAGGGGGCCCGCTCTGGGTGAAGACTCTGTGCAAGATGTGGGGATG	1860
Db	1801	TCTCTCCGACGAGAGGGGGCCCGCTCTGGGTGAAGACTCTGTGCAAGATGTGGGGATG	1860
QY	1861	CAGCTGTGGCATGGAATGATCTGGAGAGCAAGTGTCTCAACCCGGGACCTGGATCTCTC	1920
Db	1861	CAGCTGTGGCATGGAATGATCTGGAGAGCAAGTGTCTCAACCCGGGACCTGGATCTCTC	1920
QY	1921	GGAATCTGCTGTGACAGAGAAAGATCTCTGAAGATCAATGTACTTTGGATGTCCGAG	1980
Db	1921	GGAATCTGCTGTGACAGAGAAAGATCTCTGAAGATCAATGTACTTTGGATGTCCGAG	1980
QY	1981	AGGAAGCCGATGGGGTCTATGCAAGCTTCAGGGGGCTCACAAGATCCCGGTAAAGTGA	2040
Db	1981	AGGAAGCCGATGGGGTCTATGCAAGCTTCAGGGGGCTCACAAGATCCCGGTAAAGTGA	2040
QY	2041	CCGCACTGAGAGCCCTTAACTACAGGCGCTACTCTTCGAAAGGCAAGTGTGAGACTTTG	2100
Db	2041	CCGCACTGAGAGCCCTTAACTACAGGCGCTACTCTTCGAAAGGCAAGTGTGAGACTTTG	2100
QY	2101	GCATCTTGTCTCTGGGAGACTTACAGCTGTGGGGCTCTCCCTTATCCCACTCAGCAATC	2160
Db	2101	GCATCTTGTCTCTGGGAGACTTACAGCTGTGGGGCTCTCCCTTATCCCACTCAGCAATC	2160
QY	2161	AGCAGACACGGGAATTTGTGGAAGAAAGGGGGCGCTGTGCGCTTCCCAAGCTGTCTCTG	2220
Db	2161	AGCAGACACGGGAATTTGTGGAAGAAAGGGGGCGCTGTGCGCTTCCCAAGCTGTCTCTG	2220
QY	2221	ATGCGCTGTTCAGGCTCATGGAAGCAAGTGTGGGCTTATGAGCTGTGGCAACGGGCCAGCT	2280
Db	2221	ATGCGCTGTTCAGGCTCATGGAAGCAAGTGTGGGCTTATGAGCTGTGGCAACGGGCCAGCT	2280
QY	2281	TCAGACCACTATTCACAGAGACTGAGAGCAATCCGAAAGCGGCAATCGGTGAGAGCTGGGAC	2340
Db	2281	TCAGACCACTATTCACAGAGACTGAGAGCAATCCGAAAGCGGCAATCGGTGAGAGCTGGGAC	2340

[illegible]

CC	XX	ADAM28580
CC	XX	ID ADAM28580 standard; cDNA; 2674 BP.
CC	XX	AC ADAM28580;
CC	XX	DT 17-JUN-2004 (first entry)
CC	XX	DE Human proto-oncogene tyrosine kinase cDNA.
CC	XX	KM sa; gene; gene therapy; human; proto-oncogene; tyrosine kinase; cancer
CC	XX	OS Homo sapiens.
CC	XX	FH Key
CC	XX	FT 5'UTR
CC	XX	Location/Qualifiers
CC	XX	1..71
CC	XX	/tag= a
CC	XX	CDS
CC	XX	72..2330
CC	XX	/tag= b
CC	XX	/product= "Proto-oncogene tyrosine kinase"
CC	XX	FT 3'UTR
CC	XX	2331..2674
CC	XX	/tag= c
CC	XX	US2004063130-A1.
CC	XX	PN 01-APR-2004.
CC	XX	PD 12-SEP-2003; 2003US-00660763.
CC	XX	PF 27-MAR-2001; 2001US-00817180.
CC	XX	PR 06-DEC-2001; 2001US-00003295.
CC	XX	PA (APPL-) APPLERA CORP.
CC	XX	Gan W, Ye J, Di Francesco V, Beasley EM;
CC	XX	WPI; 2004-282461/26.
CC	XX	DR P-PDB; ADM28581.
CC	XX	PT New human kinase peptides, useful for preparing a composition for
CC	XX	treating a disease or condition mediated by human kinases.
CC	XX	PS Claim 4; SEQ ID NO 1; 53bp; English.
CC	XX	The invention relates to a new isolated human proto-oncogene tyrosine
CC	XX	kinase peptide or its allelic variant, orthologue or fragment. The
CC	XX	peptide is useful for preparing a composition for treating a disease or
CC	XX	condition mediated by a human kinase protein e.g. cancer. The present

CC	sequence	represents the human proto-oncogene tyrosine kinase cDNA.
XX	Sequence	2674 bp; 662 A; 767 C; 810 G; 435 T; 0 U; 0 Other;
SQL	Query Match	100.0%; Score 2674; DB 12; Length 2674;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 2674; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1	TCGCGGGTCGCGACCCGGGCTGAGTCGGTCCGAGGCCGTCCGAGAGCAGTGCCTGCGTGC 60
DB	1	TCCGGGGGTCGCGACCCGGGCTGAGTCGGTCCGAGGCCGTCCGAGAGCAGTGCCTGCGTGC 60
QY	61	GGAACAGACATATGGGGCTCTCTTCTGAGCTGAGACCCCGGAGGGGCGACGGGGTCTCTGC 120
DB	61	GGAACAGACATATGGGGCTCTCTTCTGAGCTGAGACCCCGGAGGGGCGACGGGGTCTCTGC 120
QY	121	AGCAAAATGCAAGAGAGCCGAGCTTGGTCTACTGAGGCGCATATGAAATGTGATGGCCAGC 180
DB	121	AGCAAAATGCAAGAGAGCCGAGCTTGGTCTACTGAGGCGCATATGAAATGTGATGGCCAGC 180
QY	181	GGGTCAAGATGACAGGAGATGACAGGAGCTGCTTCCACATGTCCCTCAGAGACAGTG 240
DB	181	GGGTCAAGATGACAGGAGATGACAGGAGCTGCTTCCACATGTCCCTCAGAGACAGTG 240
QY	241	GGGGCCAGAGCGGGGGCCATCAGCCCTTCACAGCCCATCATTCAGTCTTGGGCTGAGATCA 300
DB	241	GGGGCCAGAGCGGGGGCCATCAGCCCTTCACAGCCCATCATTCAGTCTTGGGCTGAGATCA 300
QY	301	CCAGCCAAATCTGAGGGGCTGAGCGGCTTGGCTGCGGCGAGCAGCAGAGGATCTGAATCTAG 360
DB	301	CCAGCCAAATCTGAGGGGCTTGAAGCGGCTTGGCTGCGGCGAGCAGCAGAGGATCTGAATCTAG 360
QY	361	GGCCCTTGAGCGAGACTGAGCCTGCTCATCCGCGAAACGGGACAGCTTCCGAAAGCTTACA 420
DB	361	GGCCCTTGAGCGAGACTGAGCCTGCTCATCCGCGAAACGGGACAGCTTCCGAAAGCTTACA 420
QY	421	GCGAGCAGTGGCAGCAGCTGTCAGCAGAGAGCTCACCAAGATCCACAGCAGAGCAATTGAGA 480
DB	421	GCGAGCAGTGGCAGCAGCTGTCAGCAGAGAGCTCACCAAGATCCACAGCAGAGCAATTGAGA 480
QY	481	AGCTGAAGAGCAGATACCGAGCTGTGGACGGGACATGTGCCCAAGCCCAAGCGCAAGTACC 540
DB	481	AGCTGAAGAGCAGATACCGAGCTGTGGACGGGACATGTGCCCAAGCCCAAGCGCAAGTACC 540
QY	541	AGGAGGCGCAGCAAAAGACAGAGACCGGTGACAAAGCCCAAGGACAATATGTGCGGACCTGT 600
DB	541	AGGAGGCGCAGCAAAAGACAGAGACCGGTGACAAAGCCCAAGGACAATATGTGCGGACCTGT 600
QY	601	GGAAGCTTTTGTCTCACCAACCGCTATGTGCTGGCGGTGCGGGCTGCGGACGTTACACC 660
DB	601	GGAAGCTTTTGTCTCACCAACCGCTATGTGCTGGCGGTGCGGGCTGCGGACGTTACACC 660
QY	661	ACGAGCACACCAACACACTCTCTGTGCGCGGCTGCTGCGGTGCTGACAGAGCCTGACAG 720
DB	661	ACGAGCACACCAACACACTCTCTGTGCGCGGCTGCTGCGGTGCTGACAGAGCCTGACAG 720
QY	721	AGGAGATGAGCTTGCATCCTGGAAGAGATCTGTGACAGAAATCTTGAAGATTAGCAGCTGG 780
DB	721	AGGAGATGAGCTTGCATCCTGGAAGAGATCTGTGACAGAAATCTTGAAGATTAGCAGCTGG 780
QY	781	TGCAGAGATGAGGTGTGCGCATTCACCCGGGAGATGTGTGAGCTGCTGCCGATCCAGC 840
DB	781	TGCAGAGATGAGGTGTGCGCATTCACCCGGGAGATGTGTGAGCTGCTGCCGATCCAGC 840
QY	841	CTGAGGCTGAGTATCCAAAGGCTTCTGTGGAACATATGGGTCCGACCTGAGGTCCCACTT 900
DB	841	CTGAGGCTGAGTATCCAAAGGCTTCTGTGGAACATATGGGTCCGACCTGAGGTCCCACTT 900
QY	901	GTCACAGCTTGCATGATGATCACTGCTTTGAGAGGGGTGAACCGCTGTGAGCTCTGGGAGCTCC 960
DB	901	GTCACAGCTTGCATGATGATCACTGCTTTGAGAGGGGTGAACCGCTGTGAGCTCTGGGAGCTCC 960
QY	961	AGCTGAACGAGCTGATCTGTGAGAGCGGTGACGACAGCTGATCCTGATGACAGATGAGC 1020

DB 961 AGCTGACGAGCTGATGAGAGGCTGACGACGCTGACCTGATGACAGATGAGC 1020
QY 1021 TGGCTGTGGCCACCGAGATGATGTTTCAAGCGCGCAGAGATGATTTGACGACCTGCAACAG 1080
DB 1021 TGGCTGTGGCCACCGAGATGATGTTTCAAGCGCGCAGAGATGATTTGACGACCTGCAACAG 1080
QY 1081 AGCTCGGAAATGAAGAGGAAACACCCACCCCGGAGGCGGTGACGTGCTGGGCAAGA 1140
DB 1081 AGCTCGGAAATGAAGAGGAAACACCCACCCCGGAGGCGGTGACGTGCTGGGCAAGA 1140
QY 1141 GGCAGTGTGCAAGAGACTGCAAGGCGCTGACGAGTACGCTGTGACGACGAGCCAGC 1200
DB 1141 GGCAGTGTGCAAGAGACTGCAAGGCGCTGACGAGTACGCTGTGACGACGAGCCAGC 1200
QY 1201 TGCAGGCCACGACGAGAGTTGCTGACAGACCAAGCTGAGCACTTGAGCCCGCGAGCCCC 1260
DB 1201 TGCAGGCCACGACGAGAGTTGCTGACAGACCAAGCTGAGCACTTGAGCCCGCGAGCCCC 1260
QY 1261 CGCTGTGCTGCTCTGCAAGATGACCGCACTCCACGTCGTCGTCGAGAGAGAGCAG 1320
DB 1261 CGCTGTGCTGCTCTGCAAGATGACCGCACTCCACGTCGTCGTCGAGAGAGAGCAG 1320
QY 1321 AGGGGGGAGAGACCCACGCTGAGATCCTTAAAGACCACTCTCAGGAATCTTCGCGC 1380
DB 1321 AGGGGGGAGAGACCCACGCTGAGATCCTTAAAGACCACTCTCAGGAATCTTCGCGC 1380
QY 1381 CCAAGTTCTCGAACCTGTACCGACGTGAAGAGGAGAGGCTTCTCTAGCAATCTCTTCTCA 1440
DB 1381 CCAAGTTCTCGAACCTGTACCGACGTGAAGAGGAGAGGCTTCTCTAGCAATCTCTTCTCA 1440
QY 1441 TGCACCACTTACGACACCCGACGACCCCTCACCAAGAGAGTGTGTCCTGACCA 1500
DB 1441 TGCACCACTTACGACACCCGACGACCCCTCACCAAGAGAGTGTGTCCTGACCA 1500
QY 1501 GGGCTGTGCCCAGAGCAAGTGGGTCTGAACCATGAGGACCTGTGTTGGGTGACAGA 1560
DB 1501 GGGCTGTGCCCAGAGCAAGTGGGTCTGAACCATGAGGACCTGTGTTGGGTGACAGA 1560
QY 1561 TTGGACGGGAGAACTTTGGCGAAAGTGTTCAGCGGACGCTGCGAGCCGACCAACCTCTGG 1620
DB 1561 TTGGACGGGAGAACTTTGGCGAAAGTGTTCAGCGGACGCTGCGAGCCGACCAACCTCTGG 1620
QY 1621 TGGCGGTGAAGTCTTTCAGAGACGCTCCACCTGACCTCAAGGCGCAAGTTCTACAG 1680
DB 1621 TGGCGGTGAAGTCTTTCAGAGACGCTCCACCTGACCTCAAGGCGCAAGTTCTACAG 1680
QY 1681 AAGCGAGATCTCGAAGCAAGTACAGCCACCCCAACATCGTGCCTCATTTGATGTCTGCA 1740
DB 1681 AAGCGAGATCTCGAAGCAAGTACAGCCACCCCAACATCGTGCCTCATTTGATGTCTGCA 1740
QY 1741 CCCAGAAAGCCCATCTACATGTGTCATGAGACTTTGTCAGGGGCGGACCTTCTGACCT 1800
DB 1741 CCCAGAAAGCCCATCTACATGTGTCATGAGACTTTGTCAGGGGCGGACCTTCTGACCT 1800
QY 1801 TCTCTCGGACGAGGGGGGCGCGCTCGGGTGAAGACTGCTGCGAGATGATGATGAGATG 1860
DB 1801 TCTCTCGGACGAGGGGGGCGCGCTCGGGTGAAGACTGCTGCGAGATGATGATGAGATG 1860
QY 1861 CAGCTCTGCGATGAGTACTTGAAGAGCAAGTGTCTGACCAACCGGAGACTTGGCTGCTC 1920
DB 1861 CAGCTCTGCGATGAGTACTTGAAGAGCAAGTGTCTGACCAACCGGAGACTTGGCTGCTC 1920
QY 1921 GGAAGTCTGCTGTGACAGAGAAATGCTCTGAAGATCAAGTGAATCTTGGGATGCTCCGAG 1980
DB 1921 GGAAGTCTGCTGTGACAGAGAAATGCTCTGAAGATCAAGTGAATCTTGGGATGCTCCGAG 1980
QY 1981 AGGAAGCCGATGGGCTGTATGACGCTTCAAGGGGCGCTCAGACAAAGTCCCGTGAAGTGA 2040
DB 1981 AGGAAGCCGATGGGCTGTATGACGCTTCAAGGGGCGCTCAGACAAAGTCCCGTGAAGTGA 2040
QY 2041 CCGCACCTGAGGCGCTTAACTTACGCGCGCTACTCTCCGAAAGCAGAGTGTGAGCTTTG 2100

DB 2041 CCGCACCTGAGGCGCTTAACTTACGCGCGCTACTCTCCGAAAGCAGCTGTGAGCTTTG 2100
QY 2101 GCAATCTTGTCTGAGAGACTTTCACGCTGGGGGCGTCCCGCATGCCAATCCAGCAATC 2160
DB 2101 GCAATCTTGTCTGAGAGACTTTCACGCTGGGGGCGTCCCGCATGCCAATCCAGCAATC 2160
QY 2161 AGCAGACACGGGAGTGTGTGAGAGAGGGGCGCTGTGCTGCTGCCAGAGCTGTGCTG 2220
DB 2161 AGCAGACACGGGAGTGTGTGAGAGAGGGGCGCTGTGCTGCTGCCAGAGCTGTGCTG 2220
QY 2221 ATGCGGTGTCAGGCTCATGAGACAGTGTGGGCTATGAGCTTGGGCGGCGCCAGCT 2280
DB 2221 ATGCGGTGTCAGGCTCATGAGACAGTGTGGGCTATGAGCTTGGGCGGCGCCAGCT 2280
QY 2281 TCAGACCAATCTACAGAGAGCTGACAGATCCGAAGCGGATGATGATGAGCTGGGACC 2340
DB 2281 TCAGACCAATCTACAGAGAGCTGACAGATCCGAAGCGGATGATGATGAGCTGGGACC 2340
QY 2341 CCTTCTCAAGCTGTGCTGCTGAGGCTTGAAGCTGAGCTCTGAGCTGCAAGCTCA 2400
DB 2341 CCTTCTCAAGCTGTGCTGCTGAGGCTTGAAGCTGAGCTCTGAGCTGCAAGCTCA 2400
QY 2401 TATGCTGACAGCTTTCACAGTCTTGAACCTCTGCGCACAGATCCACTGCGGACAG 2460
DB 2401 TATGCTGACAGCTTTCACAGTCTTGAACCTCTGCGCACAGATCCACTGCGGACAG 2460
QY 2461 ATGACGCGCGCTGCTCTGCTGTGTCCTGCTGCTGCGGCGGCTTCTTCCGGGACAG 2520
DB 2461 ATGACGCGCGCTGCTCTGCTGTGTCCTGCTGCTGCGGCGGCTTCTTCCGGGACAG 2520
QY 2521 AACAAATTAACCACTGTGCGCCACTGAAAAA 2580
DB 2521 AACAAATTAACCACTGTGCGCCACTGAAAAA 2580
QY 2581 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2640
DB 2581 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2640
QY 2641 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674
DB 2641 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674

RESULT 4
ADK71888
ID ADK71888 standard; cDNA; 2599 BP.
XX
AC ADK71888;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human kinase and phosphatase KPP-6 cDNA.
XX
XX human; kinase; phosphatase; KPP; cardiovascular; antiarteriosclerotic;
XX hypotensive; vasodilator; antiinflammatory; antihypertensive; anti-HIV;
XX antiallergic; antiaesthetic; immunosuppressive; antidiabetic;
XX dermatological; antidiabetic; nephrotoxic; antitumor; gastrointestinal;
XX neuroprotective; osteoporosis; antitoxic; anticonvulsant; hepatotoxic;
XX antirheumatic; antiparkinsonian; nootropic; anticonvulsant; hepatotoxic;
XX antileptotic; haemostatic; cytosolic; antileptotic; antiparasitic;
XX antileptotic; haemostatic; cytosolic; antileptotic; antiparasitic;
XX cardiovascular disease; immune system; neurological; growth; development;
XX cell proliferation; viral; bacterial; fungal; parasitic; protozoan;
XX helminthic; infection; transgenic; gene therapy; ss; gene;
XX single nucleotide polymorphism; SNP.
OS Homo sapiens.
XX
XX
FH key Location/Qualifiers
FT variation replace(109,T)
FT /tag= a
FT /label= Single nucleotide polymorphism (SNP)
XX

PN W02004018641-A2.
 XX
 PD 04-MAR-2004.
 XX
 PF 25-AUG-2003; 2003WO-US026635.
 XX
 PR 26-AUG-2002; 2002US-0406172P.
 PR 25-SEP-2002; 2002US-0413910P.
 PR 27-SEP-2002; 2002US-0414296P.
 PR 11-OCT-2002; 2002US-0417821P.
 XX
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Baughn MR, Richardson TW, Marquis JP, Swarnakar A, Tang YT,
 PI Becha SD, Emeiling BM, Jin P, Wilson AD, Yue H, Gietzen KU,
 PI Chang H, Yang YG, Lee SY, Khare R, Elliott VS, Hafalia AA,
 PI Chavala NK, Ramkumar J, Gurusajan R, Tribouley CM, Chien D, Tran UK,
 PI Murage J;
 XX
 XX MPI: 2004-226830/21.
 DR P-PSDB; ADK71829.
 DR
 XX
 XX New human kinases and phosphatases, useful for diagnosing, treating or
 PT preventing atherosclerosis, hypertension, AIDS, allergy, multiple
 PT sclerosis, osteoarthritis, Alzheimer's disease, Crohn's disease, cancer
 PT or hepatitis.
 XX
 PS Claim 5; SEQ ID NO 65; 347bp; English.
 XX
 CC The invention relates to a novel isolated polypeptide which is a human
 CC kinase and phosphatase (KRP). The polypeptide of the invention
 CC demonstrates cardiovascular, antiarteriosclerotic, hypotensive,
 CC vasotropic, antiinflammatory, antidiabetic, anti-HIV, antiallergic,
 CC antiaesthetic, immunosuppressive, antihypertensive, dermatological,
 CC antidiabetic, nephroprotective, antispasmodic, gastrointestinal, neuroprotective,
 CC osteoparalytic, antirheumatic, uropathic, ophthalmological, antihemiparetic,
 CC antiparkinsonian, nociceptive, anticonvulsant, hepatotropic, antiparasitic,
 CC haemostatic, cytostatic, antilipemic, antiparasitic, antihelminthic,
 CC antibacterial, virucide, protozoacide and fungicide activities. The
 CC kinase and phosphatase (KRP) polynucleotides, polypeptides, agonists and
 CC antagonists may be useful for diagnosing, treating or preventing
 CC disorders such as cardiovascular diseases, immune system disorders,
 CC neurological disorders, disorders affecting growth and development, cell
 CC proliferative disorders and viral, bacterial, fungal, parasitic,
 CC protozoan or helminthic infections. Furthermore, the molecules of the
 CC invention may be useful for creating transgenic animals to model human
 CC disease and during gene therapy. The current sequence is that of a human
 CC KRP cDNA of the invention.
 CC
 XX
 SQ Sequence 2599 BP, 541 A, 786 C, 828 G, 444 T, 0 U, 0 Other;
 Query Match 93.1%; Score 2490.2; DB 12; Length 2599;
 Best Local Similarity 98.5%; Pred. No. 0;
 Matches 2537; Conservative 0; Mismatches 3; Indels 35; Gaps 1;

Db 265 GGGCCAGAGCCGGGGCCATGAGCCCTGACAGCCCCATCACTGCTCGGGCTGAGATCAC 324
 QY 302 CAGCCAAATTAGAGGCTTGAAGCCCTTGTCTGGGGCAGCAGCAGAGATCTGAATCAGG 361
 Db 325 CAGCCAACTGAGGGGCTGAGCCGCTTGTCTGGGGCAGCAGCAGAGATCTGAATCAGG 384
 QY 362 GCGCCCTGAGCAAGCTGAGCTGCTCATCCGGGAAAGGGCAGAGCTTCCGAAAGCTACAG 421
 Db 385 GCGCCCTGAGCAAGCTGAGCTGCTCATCCGGGAAAGGGCAGAGCTTCCGAAAGCTACAG 444
 QY 422 CGAGCAGTGGCAGAGCTTGCAGCAGAGAGCTCAACAAAGCCACAGCCAGAGACTTGAAGA 481
 Db 445 CGAGCAGTGGCAGAGCTTGCAGCAGAGAGCTCAACAAAGCCACAGCCAGAGACTTGAAGA 504
 QY 482 GCTGAAGAGCCAGTACCGAGCTTGTGGCAGGAGCAGTGTCCCAAGCCAGAGAGTACCA 541
 Db 505 GCTGAAGAGCCAGTACCGAGCTTGTGGCAGGAGCAGTGTCCCAAGCCAGAGAGTACCA 564
 QY 542 GGAAGCCAGCAAGAAAGCAAGAGACCGTGAACAAAGCCAAAGTATGTGGAGCCCTGTG 601
 Db 565 GGAAGCCAGCAAGAAAGCAAGAGACCGTGAACAAAGCCAAAGTATGTGGAGCCCTGTG 624
 QY 602 GAAGCTCTTGTGTCACCAACACCGCTATGTGTGGCGCTGCGGGCTGCGCAGCTACACCA 661
 Db 625 GAAGCTCTTGTGTCACCAACACCGCTATGTGTGGCGCTGCGGGCTGCGCAGCTACACCA 684
 QY 662 CAGACACACACACAGCTCTGTGCTGCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
 Db 685 CAGACACACACACAGCTCTGTGCTGCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 744
 QY 722 GGAAGTGGCTTGCATCTGAAGAGATCTGACAGAAATCTGAGATTAAGACCTGGGT 781
 Db 745 GGAAGTGGCTTGCATCTGAAGAGATCTGACAGAAATCTGAGATTAAGACCTGGGT 804
 QY 782 GAGAGTGAAGTGTGGCCATTACCGGAGAGATGTGCTGAGCTGTCCCGCATCCAGCC 841
 Db 805 GAGAGTGAAGTGTGGCCATTACCGGAGAGATGTGCTGAGCTGTCCCGCATCCAGCC 864
 QY 842 TGAAGTGAATCAAGGCTTCTGTGACAGATATGAGTGTGCTGAGCTGTCCCGCATCCAG 901
 Db 865 TGAAGTGAATCAAGGCTTCTGTGACAGATATGAGTGTGCTGAGCTGTCCCGCATCCAG 924
 QY 902 TGTCAAGTGAATCAAGGCTTCTGTGACAGATATGAGTGTGCTGAGCTGTCCCGCATCCAG 961
 Db 925 TGTCAAGTGAATCAAGGCTTCTGTGACAGATATGAGTGTGCTGAGCTGTCCCGCATCCAG 984
 QY 962 GCTGAAGCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1021
 Db 985 GCTGAAGCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1044
 QY 1022 GGTGTGTGACCCAGAGATGTGTTCAGGCGGAGAGATGTGTTCAGGCGGAGAGATGTGTTCAG 1081
 Db 1045 GGTGTGTGACCCAGAGATGTGTTCAGGCGGAGAGATGTGTTCAGGCGGAGAGATGTGTTCAG 1104
 QY 1082 GCTTCGGAGTGAAG 1141
 Db 1105 GCTTCGGAGTGAAG 1164
 QY 1142 GCAAGTGTGCAAG 1201
 Db 1165 GCAAGTGTGCAAG 1224
 QY 1202 GCAAGTGTGCAAG 1261
 Db 1225 GCAAGTGTGCAAG 1284
 QY 1262 GCTGTGTGCTTCGAG 1306
 Db 1285 GCTGTGTGCTTCGAG 1344
 QY 1307 GCTGTGTGCTTCGAG 1346
 Db 1345 GACGTCTGTCCCTGAGCTCAG 1404

```
QY 1347 ATCTTTAAGAGCCACATCTCAGGAATCTTCCGCCCCCAAGTTCTCGAACCTGTACCGACCTG 1406
DB 1405 ATCTTTAAGAGCCACATCTCAGGAATCTTCCGCCCCCAAGTTCTCGAACCTGTACCGACCTG 1464
QY 1407 GAAGGGGAAGGCTTTCTAGCATCTCTTGTCTCATGACCACTTACTGACACCCAGCAG 1466
DB 1465 GAAGGGGAAGGCTTTCTAGCATCTCTTGTCTCATGACCACTTACTGACACCCAGCAG 1524
QY 1467 CCCCTTACCAAGAAAGTGTGTGTCTGTGCAACAGGGCTGTGTGCCCAAGAACAGTGGGTG 1526
DB 1525 CCCCTTACCAAGAAAGTGTGTGTCTGTGCAACAGGGCTGTGTGCCCAAGAACAGTGGGTG 1584
QY 1527 CTGAACCATGAGGACCTGTGTGTGTGTGAGACAGATTGACAGGGGGAACCTTTGGCGAAGTG 1586
DB 1585 CTGAACCATGAGGACCTGTGTGTGTGTGAGACAGATTGACAGGGGGAACCTTTGGCGAAGTG 1644
QY 1587 TTTCAGCGGACGCTGTGCGAGCCGACACACCTGTGTGGCGGTGAAAGTCTTGTGAGAGACG 1646
DB 1645 TTTCAGCGGACGCTGTGCGAGCCGACACACCTGTGTGGCGGTGAAAGTCTTGTGAGAGACG 1704
QY 1647 CTCCCACTGACCTTCAAGGCGCAAGTTTCTACAGAAAGCCAGAGATCTTGAAGCAGTACAGC 1706
DB 1705 CTCCCACTGACCTTCAAGGCGCAAGTTTCTACAGAAAGCCAGAGATCTTGAAGCAGTACAGC 1764
QY 1707 CACCCCAACATGATGAGTGTCTCATTTGGTGTCTGACCCAGAAAGAGCCCATCTACATGCTC 1766
DB 1765 CACCCCAACATGAGTGTGTCTCATTTGGTGTCTGACCCAGAAAGAGCCCATCTACATGCTC 1824
QY 1767 ATGAGACTTGTGAGGGGGGGGCGACTTCTGACCTTCTCTCGACCGAGGGGGCGGCGCTG 1826
DB 1825 ATGAGACTTGTGAGGGGGGGGCGACTTCTGACCTTCTCTCGACCGAGGGGGCGGCGCTG 1884
QY 1827 CGGGTAAAGACTTGTGAGAGATGTGTGGGGGAATGACAGCTGTGTGGCATGTGAGTACCTGAG 1886
DB 1885 CGGGTAAAGACTTGTGAGAGATGTGTGGGGGAATGACAGCTGTGTGGCATGTGAGTACCTGAG 1944
QY 1887 AGCAAGTGTGATCTACACCGGAGACTGTGCTGTGCGAACTGTGTGACAGAGAAAT 1946
DB 1945 AGCAAGTGTGATCTACACCGGAGACTGTGCTGTGCGAACTGTGTGACAGAGAAAT 2004
QY 1947 GTCTTGAAGATCAGTACTTGTGGAGTGTCCGAGAGAAAGCCGATGGGGTCTATGACGCG 2006
DB 2005 GTCTTGAAGATCAGTACTTGTGGAGTGTCCGAGAGAAAGCCGATGGGGTCTATGACGCG 2064
QY 2007 TCGAGGGGGCTCAGACAAAGTCCCCGTAAGTGAACCGCACCTTAAAGCCCTTAACGCGC 2066
DB 2065 TCGAGGGGGCTCAGACAAAGTCCCCGTAAGTGAACCGCACCTTAAAGCCCTTAACGCGC 2124
QY 2067 CGCTACTCCCTCCGAAAGCGAGCTGTGAGCTTGGCATCTTGTCTGTGGAGACCTTCAGC 2126
DB 2125 CGCTACTCCCTCCGAAAGCGAGCTGTGAGCTTGGCATCTTGTCTGTGGAGACCTTCAGC 2184
QY 2127 CTGGGGGGCTCCCTCATATCCCAACTTACAGAAATCAGACAGACAGGGAGTTTGTGAGAAAG 2186
DB 2185 CTGGGGGGCTCCCTCATATCCCAACTTACAGAAATCAGACAGACAGGGAGTTTGTGAGAAAG 2244
QY 2187 GGGGGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2246
DB 2245 GGGGGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2304
QY 2247 TGTGTGGGCTATAGAGCTGTGGGAGCGGAGCCAGCTTACAGACCACTTACAGAGAGCTGACG 2306
DB 2305 TGTGTGGGCTATAGAGCTGTGGGAGCGGAGCCAGCTTACAGACCACTTACAGAGAGCTGACG 2364
QY 2307 AGCATCCGAAAGCGGACATCGGTGAGAGCGGAGCCCTTCTCAAGCTGTGTGCTGTGCA 2366
DB 2365 AGCATCCGAAAGCGGACATCGGTGAGAGCGGAGCCCTTCTCAAGCTGTGTGCTGTGCA 2424
QY 2367 GGGCTTGTGTGAGCTCTCAGCGGCTTCACTATGTGCTGACAGCTCTTCAAGCTGTGCA 2426
DB 2425 GGGCTTGTGTGAGCTCTCAGCGGCTTCACTATGTGCTGACAGCTCTTCAAGCTGTGCA 2484
```

```
QY 2427 GACTCTGTGCAACAGATCTCAGAGTGTGCGGAGATGACAGCGCGCTGTCTGTGTGTC 2486
DB 2485 GACTCTGTGCAACAGATCTCAGAGTGTGCGGAGATGACAGCGCGCTGTCTGTGTGTC 2544
QY 2487 CCTGTGTGTGCGGAGCTTCTTCTTCCGAGGAGAAACATTAACCATTTGTGCCC 2541
DB 2545 CCTGTGTGTGCGGAGCTTCTTCTTCCGAGGAGAAACATTAACCATTTGTGCCC 2599

RESULT 5
AAS95001
ID AAS95001 standard; DNA; 2889 BP.
XX
AC AAS95001;
XX
DT 14-FEB-2002 (first entry)
XX
DE Human DNA sequence #256 expressed during foam cell differentiation.
XX
KW Human; foam cell differentiation; atherosclerosis; cerebral stroke;
XX cardiovascular disorder; coronary artery disease; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200177389-A2.
XX
PD 18-OCT-2001.
XX
PF 04-APR-2001; 2001WO-US011128.
XX
PR 05-APR-2000; 2000US-0195106F.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Shiffman D, Somogyi R, Lawn R, Selthamer J, Porter GJ, Mikita T;
PI Tai J;
XX
DR WPI; 2002-010925/01.
XX
PT Composition useful for diagnosis of conditions, disorders or diseases
PT associated with atherosclerosis, comprises several polynucleotides that
PT are differentially expressed in foam cell development.
XX
PS Claim 1; Page 298-299; 315pp; English.
XX
CC The present invention relates to the isolation of human polynucleotide
CC sequences that are differentially expressed during foam cell
CC differentiation. The polynucleotide sequences of the invention or a
CC composition comprising these polynucleotides are useful as a high
CC throughput method for detecting altered expression of one or more
CC polynucleotides in a sample. The polynucleotides can be used in the
CC diagnosis of disorders associated with foam cell development such as
CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC coronary artery disease. The polynucleotide sequences can also be used as
CC PCR primers and probes. The polynucleotides of the invention are also
CC useful in gene therapy. AAS94746-AAS95021 represent the human
CC polynucleotide sequences of the invention which are differentially
CC expressed during foam cell differentiation
XX
SQ Sequence 2889 BP; 600 A; 873 C; 929 G; 487 T; 0 U; 0 Other;

Query Match 87.0%; Score 2326.4; DB 6; Length 2889;
Best Local Similarity 92.3%; Pred. No. 6; 8e-305;
Matches 2558; Conservative 0; Mismatches 1; Indels 21; Gaps 2;

QY 2 CCGGGGCTCGCACCCGGGCTGAGTGTGCGAGAGCCGTCCAGAGACAGTGTCCGTGCG 61
DB 118 CCGGGGCTCGCACCCGGGCTGAGTGTGCGAGAGCCGTCCAGAGACAGTGTCCGTGCG 177
QY 62 GAACAGACTTATGGGCTTCTTCTGTGAGCTGTGTGAGAGCCCGCAGGGGCGAGGGGTCTGCA 121
DB 178 GAACAGACTTATGGGCTTCTTCTGTGAGCTGTGTGAGAGCCCGCAGGGGCGAGGGGTCTGCA 237
```

122 GCAATGACGAGGAGCCGAGCTTGCTACTAGAGGGGCGATGAGAAAGTGGATGGCCGAGG 181
123 GCAATGACGAGGAGCCGAGCTTGCTACTAGAGGGGCGATGAGAAAGTGGATGGCCGAGG 297
182 GGTCAAGAGTGA CAGGAGATATG CAGAGACTGCTTCA CCAATGTCCCTG CAGAGCAGTGG 241
298 GGTCAAGAGTGA CAGGAGATATG CAGAGACTGCTTCA CCAATGTCCCTG CAGAGCAGTGG 357
242 GGGCCAGAGCCGGGCGCATAGCCCTGACAGCCCATAGTCAAGTCTCTGGGCTGAGATCAC 301
358 GGGCCAGAGCCGGGCGCATAGCCCTGACAGCCCATAGTCAAGTCTCTGGGCTGAGATCAC 417
302 GAGCCAAACTGA -GGGCTGAGGCGGCTTGCTGCGGAGACCGAGAGAGATCTGAACTCAG 360
418 GAGCCAAACTGAGGGGCTGAGCGGCTTGCTGCGGAGACCGAGAGAGATCTGAACTCAG 477
361 GGGCCCTGACGAGCTGAGCTGAGCTTCTCATCCGGGAA CCGGAGCAGTTCGCAAGACCTTAC 420
478 GGGCCCTGACGAGCTGAGCGCTTCTCATCCGGGAA CCGGAGCAGCTTCCGAGAACTTACA 537
421 GCGAGCAGTGGCAGCAGCTGACAGCAGAGCTCACCAAGACCCACAGCCAGACATTGAGA 480
538 GCGAGCAGTGGCAGCAGCTGACAGCAGAGCTCACCAAGACCCACAGCCAGACATTGAGA 597
481 AGCTGAAGAGCCAGTACCGAGCTTGCGACCGGACAGTCCCAAGCCCAAGCCGAAAGTACC 540
598 AGCTGAAGAGCCAGTACCGAGCTTGCGACCGGACAGTCCCAAGCCCAAGCCGAAAGTACC 657
541 AGGAGCCGACAAAGACAAAGACCGTGA CAGGCGCAAGGACAAAGTATGTCGAGCAGCTGT 600
658 AGGAGCCGACAAAGACAAAGACCGTGA CAGGCGCAAGGACAAAGTATGTCGAGCAGCTGT 717
601 GGAAGCTCTTTGCTCA CCAACACCGTATGTCGTGGGCGCTGCGGAGCTGCGAGCTACAC 660
718 GGAAGCTCTTTGCTCA CCAACACCGTATGTCGTGGGCGCTGCGGAGCTGCGAGCTACAC 777
661 ACCAGACACACACAGCTCTGCTGCGGCGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGAG 720
778 ACCAGACACACACAGCTCTGCTGCGGCGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGAG 837
721 AGGAGATGAGCTTGCATCTGAGAGAGATCTGAGAGAAATCTGAGAGATTAGACAGCTGG 780
838 AGGAGATGAGCTTGCATCTGAGAGAGATCTGAGAGAAATCTGAGAGATTAGACAGCTGG 897
781 TGCAGAGTGAAGTGTGGCCATTACCGGAGAGATGCTGCAAGTCTGCTGCCGATCCAGC 840
898 TGCAGAGTGAAGTGTGGCCATTACCGGAGAGATGCTGCAAGTCTGCTGCCGATCCAGC 957
841 CTGAGGCTGAGTACCAAGGCTTCTGCGACAGTATGGGTCCGACCTGAGACGTCCACCT 900
958 CTGAGGCTGAGTACCAAGGCTTCTGCGACAGTATGGGTCCGACCTGAGACGTCCACCT 1017
901 GTGTCACTTGCATGATGATCTGCTGAGAGAGGAGTAA CCGCTGAGAGCTGGGAGAGCTCC 960
1018 GTGTCACTTGCATGATGATCTGCTGAGAGAGGAGTAA CCGCTGAGAGCTGGGAGAGCTCC 1077
961 AGCTGAACGAGCTGACTGTGAGAGAGGCTGACAGCACCGCTGACCTGACATGACAGATGAGC 1020
1078 AGCTGAACGAGCTGACTGTGAGAGAGGCTGACAGCACCGCTGACCTGACATGACAGATGAGC 1137
1021 TGGCTGTGGCCACCGAGATGTGTTCAGGCGGACAGAGATGTGTTAAGCAGACTGCAACAG 1080
1138 TGGCTGTGGCCACCGAGATGTGTTCAGGCGGACAGAGATGTGTTAAGCAGACTGCAACAG 1197
1081 AGCTCCGAGATGAAGAGAGAACCA CCCCACCGGAGGCGGAGTGAAGCTGTGGGAGAG 1140
1198 AGCTCCGAGATGAAGAGAGAACCA CCCCACCGGAGGCGGAGTGAAGCTGTGGGAGAG 1257
1141 GGCAGAGTGTGCAAGAGACCTG CAGAGGAGTGCAGGATAGGCTGTG CAGGAGGAGG 1200
1258 GGCAGAGTGTGCAAGAGACCTG CAGAGGAGTGCAGGATAGGCTGTG CAGGAGGAGG 1317
1201 TGCAGGCCACGACGAGAGTGTG CAGACCAAGCTGAGACACTGTGGGCCCCGGGAGAGCCC 1260

1318 TGCAGGCCACGACGAGAGTGTG CAGACCAAGCTGAGACACTGTGGGCCCCGGGAGAGCCC 1377
1261 CGCTGTGTGCTCTTGCAGAGTGA CCGGCACTTCCAGTGTCTTCGAGCAGAGACGAG 1320
1378 CGCTGTGTGCTCTTGCAGAGTGA CCGGCACTTCCAGTGTCTTCGAGCAGAGACGAG 1437
1321 AGGGGGAAAGACACCCAGCTGAGAGATCTTTAAGAGCCACATCTTCAAGAAATCTTCGCC 1380
1438 AGGGGGAAAGACACCCAGCTGAGAGATCTTTAAGAGCCACATCTTCAAGAAATCTTCGCC 1497
1381 CCAAGTTCTCG----- 1391
1498 CCAAGTTCTCGCTCCCTCCACCGCTGACGCTATTCGAGAGGTGAGAGAGCCCTGCAATG 1557
1392 ----- 1391
1558 AGCAGCTGTGTACACAGGGGCGATCCGAGGG CAGAGGTGCTGAGCTGTGTGCACT 1617
1392 ----- 1391
1618 CTGGGAGCTTCTGTGTGCGGAGAGCCAGGACAGAGAGTACGTGTGTGCTGT 1677
1392 ----- 1391
1678 GGGATGTGTGCGCCCGGCACTTCATCATCTCAAGTCTTGGATTAACCTGTACCACTGAAAG 1737
1411 GGGAAAGCTTTCTTACATCTTGTGCTATGACACCACTACTGAGACACCCAGAGCCCC 1470
1738 GGGAAAGCTTTCTTACATCTTGTGCTATGACACCACTACTGAGACACCCAGAGCCCC 1797
1471 TCACCAAGAAAGTGTGTGTGTCTG CACAGGGGCTGTGCCCAAGAGCAAGTGGTCTGA 1530
1798 TCACCAAGAAAGTGTGTGTGTCTG CACAGGGGCTGTGCCCAAGAGCAAGTGGTCTGA 1857
1531 ACCATGAGGACCTGGGTGTGGGTGAGCAGATTGSA CCGGGGAACTTTGGCGAAAGTGTCA 1590
1858 ACCATGAGGACCTGGGTGTGGGTGAGCAGATTGSA CCGGGGAACTTTGGCGAAAGTGTCA 1917
1591 GCGGAGCGCTTGAGGCGGACCA CACCCTGTGTGGGAGTGAAGTCTTGTGAGAGACGCTCC 1650
1918 GCGGAGCGCTTGAGGCGGACCA CACCCTGTGTGGGAGTGAAGTCTTGTGAGAGACGCTCC 1977
1651 CACTGACCTCAAGGCGAAAGTTTCTA CAGAAAGCGAGATCTTGAAGCAGTACAGCCACC 1710
1978 CACTGACCTCAAGGCGAAAGTTTCTA CAGAAAGCGAGATCTTGAAGCAGTACAGCCACC 2037
1711 CCAACATGATGGGTCTCATTTGGTGTG CACCCAGAAAGAGCCCATCTATGTCATGG 1770
2038 CCAACATGATGGGTCTCATTTGGTGTG CACCCAGAAAGAGCCCATCTATGTCATGG 2097
1771 AGCTGTGACGAGGGGCGACTTCTG CACCTTCCGACCGAGGGGGCCCGCTGCGGG 1830
2098 AGCTGTGACGAGGGGCGACTTCTG CACCTTCCGACCGAGGGGGCCCGCTGCGGG 2157
1831 TGAAGACTTGTCTG CAGATGTGTGGGAGTGA CAGCTGTCTG CAGATGATACCTGSAAGCA 1890
2158 TGAAGACTTGTCTG CAGATGTGTGGGAGTGA CAGCTGTCTG CAGATGATACCTGSAAGCA 2217
1891 AGTGTGCAATCCACCGGGGACTGTG CAGTCCGAACTGCTGGGTGACAGAGAAATGTCC 1950
2218 AGTGTGCAATCCACCGGGGACTGTG CAGTCCGAACTGCTGGGTGACAGAGAAATGTCC 2277
1951 TGAAGATCAGTGA CTTTGGGATGTCCCGA GAGAAAGCCAGTGGGATCTATG CAGCCTCAG 2010
2278 TGAAGATCAGTGA CTTTGGGATGTCCCGA GAGAAAGCCAGTGGGATCTATG CAGCCTCAG 2237
2011 GGGGCTTCA GACAGTGTCCCGTGAAGTGA CCGCAGCTTGAAGGCTTTAATCTAGGCGGCT 2070
2238 GGGGCTTCA GACAGTGTCCCGTGAAGTGA CCGCAGCTTGAAGGCTTTAATCTAGGCGGCT 2297
2071 ACTCTCCGAAAGCGAGTGTGAGAGCTTGG CACTTGTGCTGTGGAGAGACTTCA GCGCTGG 2130

DB 2398 ACTCCTCCGAAAGCAGCTGTGGAGCTTTGGACTTGTCTGGAGACCTTCAGCCTG 2457
 QY 2131 GGGCCCTCCCTATATCCCACTCAGAAATCAGAGACAGGGAGTTGGGAGAGGGGG 2190
 DB 2458 GGGCCCTCCCTATATCCCACTCAGAAATCAGAGACAGGGAGTTGGGAGAGGGGG 2517
 QY 2191 GCGCTGTGCGCCGCGCAGAGCTGTGTCTGATGCCGTGTTCAAGGCTCATGAGACAGTGT 2250
 DB 2518 GCGCTGTGCGCCGCGCAGAGCTGTGTCTGATGCCGTGTTCAAGGCTCATGAGACAGTGT 2577
 QY 2251 GGGCCTATAGAGCTGGGCGAGCGGCCAGCTTCAGACCACTTACAGAGAGCTGACAGCA 2310
 DB 2578 GGGCCTATAGAGCTGGGCGAGCGGCCAGCTTCAGACCACTTACAGAGAGCTGACAGCA 2637
 QY 2311 TCCGAAAGGGGATGGGTGAGAGCTGGGACCCCTTCTCAAGCTGGTGGCTGACAGGC 2370
 DB 2638 TCCGAAAGGGGATGGGTGAGAGCTGGGACCCCTTCTCAAGCTGGTGGCTGACAGGC 2697
 QY 2371 TAGGTGACAGCTCTCAGCGGCTTCAGACTATATGCTGACAGCTTCACAGTCTTGAGCT 2430
 DB 2698 TAGGTGACAGCTCTCAGCGGCTTCAGACTATATGCTGACAGCTTCACAGTCTTGAGCT 2757
 QY 2431 CCTGCCACAGCATCACAAGCTGCGGCGAGAGTACAGCGCGTGTCTCTGTGTCTG 2490
 DB 2758 CCTGCCACAGCATCACAAGCTGCGGCGAGAGTACAGCGCGTGTCTCTGTGTCTG 2817
 QY 2491 CTGCTGCGAGGGGCTTCTCTCCGGGCGAAGAACATTAACCACTTGTGCCCATGAGAA 2550
 DB 2818 CTGCTGCGAGGGGCTTCTCTCCGGGCGAAGAACATTAACCACTTGTGCCCATGAGAA 2877
 QY 2551 AAAAAAAAAA 2560
 DB 2878 AAAAAAAAAA 2887
 RESULT 6
 ADR39816
 ID ADR39816 standard; cDNA; 2623 BP.
 AC ADR39816;
 XX
 XX 18-NOV-2004 (first entry)
 DE Human kinase and phosphatase KPP-43 encoding cDNA SEQ ID NO:89.
 XX
 XX human; kinase and phosphatase protein; KPP; enzyme; cytosolic;
 KW antiarteriosclerotic; anticonvulsant; neurotropic; neuroprotective;
 KW cerebroprotective; anti-HIV; anti-allergic; anti-inflammatory;
 KW chymotrypsin; gene therapy; cell proliferative disorder; cancer;
 KW atherosclerosis; neurological disorder; epilepsy; Huntington's disease;
 KW stroke; immune disorder; inflammatory disorder; AIDS; allergy;
 KW developmental disorder; Hypothyroidism; Cushing's syndrome; infection;
 KW KPP-43; single nucleotide polymorphism; SNP; gene; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 95..2389
 FT /tag= a
 FT /product= "kinase and phosphatase KPP-43"
 FT /replac= (109..C)
 FT /tag= b
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 XX
 XX WO2004074453-A2.
 XX
 XX 02-SEP-2004.
 PD 20-FEB-2004; 2004MO-US005092.
 XX
 XX 20-FEB-2003; 2003US-0449059P.
 PR 19-MAR-2003; 2003US-0456932P.
 PR 28-MAR-2003; 2003US-0458844P.
 PR

PR 09-APR-2003; 2003US-0461678P.
 PR 17-APR-2003; 2003US-0463937P.
 XX
 PA (INCYTE) INCYTE CORP.
 XX
 PI Ramkumar J, Margulis JP, Swarnakar A, Chawla NK, Tran UK,
 PI Becha SD, Lee SY, Hafalia Aja, Richardson TW, Khare R, Jiang X;
 PI Jackson AA, Yang J, Gorvad AB;
 XX
 DR MPI; 2004-635568/61.
 DR P-PSDB; ADR39770.
 XX
 XX New human kinases and phosphatases (KPP) for diagnosing, treating and
 PT preventing diseases or conditions associated with aberrant KPP expression
 PT e.g. cancer, acquired immunodeficiency syndrome, epilepsy, or infections.
 PS Claim 5; SEQ ID NO 89; 299pp; English.
 XX
 XX The present sequence encodes the human kinase and phosphatase protein
 CC (KPP), designated KPP-43. The human KPP sequences from the present
 CC invention have cytosolic, antiarteriosclerotic, anticonvulsant,
 CC neurotropic, neuroprotective, cerebroprotective, anti-HIV, anti-allergic,
 CC anti-inflammatory and chymotrypsin activities, and can be used in gene
 CC therapy. The human KPP proteins and polynucleotides can be used in
 CC diagnosing, treating and preventing diseases or conditions associated
 CC with the decreased expression or overexpression of KPP, such as cell
 CC proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
 CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
 CC allergies), and developmental (e.g. Hypothyroidism, Cushing's syndrome)
 CC disorders, or infections. They can also be used in assessing the effects
 CC of exogenous compounds on the expression of nucleic acid and amino acid
 CC sequences of KPP. The KPP or its fragments are useful in screening
 CC compounds for effectiveness as agonist or antagonist of the polypeptides,
 CC or in altering the expression of the target polynucleotide and compounds
 CC that specifically bind to or modulate the activity of the polypeptide.
 XX
 SQ Sequence 2623 BP; 541 A; 786 C; 840 G; 456 T; 0 U; 0 Other;
 Query Match 73.0%; Score 1953; DB 13; Length 2623;
 Best Local Similarity 85.8%; Pred. No. 1,5e-254;
 Matches 2360; Conservative 0; Mismatches 5; Indels 384; Gaps 2;
 QY 2 CCGGGGTCCGACCGGGCTGAGTGGTCCGAGCGCGTCCAGAGACCTGCGCG 61
 DB 25 CCGGGGTCCGACCGGGCTGAGTGGTCCGAGCGCGTCCAGAGACCTGCGCG 84
 QY 62 GAACAGCATTAAGGCTTCTTCTTGAAGTGGACGCCCCAGGGGCTCTTGA 121
 DB 85 GAACAGCATTAAGGCTTCTTCTTGAAGTGGACGCCCCAGGGGCTCTTGA 144
 QY 122 GCAAAATGAGAGGCGGAGCTTCTTCTTGAAGTGGACGAGAAATGATGGCCAGCG 181
 DB 145 GCAAAATGAGAGGCGGAGCTTCTTCTTGAAGTGGACGAGAAATGATGGCCAGCG 204
 QY 182 GGTCAAGAGTGAAGAGATATGACAGACTCTTCAACATGTCCTTGAGAGACAGTGG 241
 DB 205 GGTCAAGAGTGAAGAGATATGACAGACTCTTCAACATGTCCTTGAGAGACAGTGG 264
 QY 242 GGGCCAGAGCGGGGCTATAGAGCTTGAAGGCCCCCATCATGCTCTGGGCTGAGTAC 301
 DB 265 GGGCCAGAGCGGGGCTATAGAGCTTGAAGGCCCCCATCATGCTCTGGGCTGAGTAC 305
 QY 302 CAGCCAACTGAGGGGCTGAGCGGCTTGTGGGCGACAGCGAGAGATCTGAATCAGG 361
 DB 306 ----- 305
 QY 362 GCCCTGAGCAAGCTGAGCTGCTCATCCGGAACGGCAGAGCTTCGAAGACTTACAG 421
 DB 306 ----- 305
 QY 422 CGAGAGTGGCGAGAGCTGACAGAGAGCTCACCAAGACCCACAGGACATTGAGAA 481
 DB 306 ----- AGACCCACAGGCGAGGACATTGAGAA 330

QY	482	GCTAAGACCAAGTACCGAGCTCTTGCGACCGGACACAGTGTCCCAAGCCACAGTACCA	541
Db	331	GCTAAGAGCCAGTACCGAGCTCTTGCGACCGGACAGTGTCCCAAGCCACAGTACCA	390
QY	542	GGAGGCCAGCAAGAACAAGACCGGTGACAAAGCCAAAGACAAGTATGTGCCAGCTGTG	601
Db	391	GGAGGCCAGCAAGAACAAGACCGGTGACAAAGCCAAAGACAAGTATGTGCCAGCTGTG	450
QY	602	GAAGCTCTTTGCTCAACCAACCCTATATGTGCTGGGCTGTGGGCTGTGCACTACCA	661
Db	451	GAAGCTCTTTGCTCAACCAACCCTATATGTGCTGGGCTGTGGGCTGTGCACTACCA	510
QY	662	CCAGCACCAACCAAGGCTCTGTGTGCCCGGCTGTGGGTGATCATGTGACGACCTGCAGCA	721
Db	511	CCAGCACCAACCAATCAGCTCTGTGTGCCCGGCTGTGGGTGATCATGTGACGACCTGCAGCA	570
QY	722	GGAGATGGCTTGATCCTGAAAGAGATCTTGCAAGAAATCCTGAGATTTAGCAGCTGGT	781
Db	571	GGAGATGGCTTGATCCTGAAAGAGATCTTGCAAGAAATCCTGAGATTTAGCAGCTGGT	630
QY	782	GCAGAGTAGAGTGGTGGCATTTCACCCGGGAATGGCTGCACCTGTGCCCGCATTCACCC	841
Db	631	GCAGAGTAGAGTGGTGGCATTTCACCCGGGAATGGCTGCACCTGTGCCCGCATTCACCC	690
QY	842	TGAGGCTGAGTACCAAGGCTTCTGCGACAGTATGGGTCCGCACTGTAGTCCCACTTG	901
Db	691	TGAGGCTGAGTACCAAGGCTTCTGCGACAGTATGGGTCCGCACTGTAGTCCCACTTG	750
QY	902	TGTCACTTTGATGATGATCATCTGTGAGAGAGGTGAACCGCTGAGCTGGGGAGCTCCA	961
Db	751	TGTCACTTTGATGATGATCATCTGTGAGAGAGGTGAACCGCTGAGCTGGGGAGCTCCA	810
QY	962	GCTAAGCAGCTGACTGTGAGAGGGGTGACACACGCTGACCTCAGTGAACAATGAGCT	1021
Db	811	GCTAAGCAGCTGACTGTGAGAGGGGTGACACACGCTGACCTCAGTGAACAATGAGCT	870
QY	1022	GGCTGTGGCCACCGAGATGATGTTTCAGGCGGCAGAGATGGTTACGCACTGCACACAGA	1081
Db	871	GGCTGTGGCCACCGAGATGATGTTTCAGGCGGCAGAGATGGTTACGCACTGCACACAGA	930
QY	1082	GCTCCGGAATGAAGAAGGAACACCCACCCCGGAGGGGTGACAGCTGTGGGCAGAG	1141
Db	931	GCTCCGGAATGAAGAAGGAACACCCACCCCGGAGGGGTGACAGCTGTGGGCAGAG	990
QY	1142	GCAAGTGTGTCAGAAAGACATGCAAGGGGCTCAGGTACCGCTGTGCAAGCCAGCCAACT	1201
Db	991	GCAAGTGTGTCAGAAAGACATGCAAGGGGCTCAGGTACCGCTGTGCAAGCCAGCCAACT	1050
QY	1202	GCAGGCCCAAGAGAGTGTGCTGCAAGCCAGACTGGAGCACTTGGGCCCGGCGAGCCCC	1261
Db	1051	GCAGGCCCAAGAGAGTGTGCTGCAAGCCAGACTGGAGCACTTGGGCCCGGCGAGCCCC	1110
QY	1262	GCCGTGTGTGCTCCGAGAGATGACCCGCACTCCACAGTGTGCTCTGAGAGCAGAGCCAGA	1321
Db	1111	GCCGTGTGTGCTCCGAGAGATGACCCGCACTCCACAGTGTGCTCTGAGAGCAGAGCCAGA	1170
QY	1322	GGGGGGAAGACACCAACGCTGAGATCTTTAAGGCCACATCTTCAGGAATCTTCCGCC	1381
Db	1171	GGGGGGAAGACACCAACGCTGAGATCTTTAAGGCCACATCTTCAGGAATCTTCCGCC	1230
QY	1382	CAAGTTTCTCG-----	1391
Db	1231	CAAGTTTCTCGCTCCTCCACCGCTGCACTCATTCGGAGGTGCAAGAACCTTCGATGA	1290
QY	1392	-----	1391
Db	1291	GCAGCTGTGTAACACCGGGGCAATCCCGAGGGCAGAGGTGGCTAGCTGCTGTGCACTC	1350
QY	1392	-----	1391
Db	1351	TGGGAGCTTCTGTGTGCGGAGAGCCAGGGCAGAGAGAGTACGTGTGTGTGCTGTG	1410

QY	1392	-----AACTGTAAACGACCTGGAAAG	1411
Db	1411	GGATGTCGTCCCGGCACTTCATCATCAAGTCTTGATTAACCTGTATCCAGCTGGAAAG	1470
QY	1412	GGAAAGCTTTCCTAGCAATCTTTTGCTCATTCGACCACTACTAGGACCCAGAGCCCT	1471
Db	1471	GGAAAGCTTTCCTAGCAATCTTTTGCTCATTCGACCACTACTAGGACCCAGAGCCCT	1530
QY	1472	CACCAAGAAAGTGTGTGTGTCCTGACAGGGCTGTGCCCAAGACAAAGTGGGTGAA	1531
Db	1531	CACCAAGAAAGTGTGTGTGTCCTGACAGGGCTGTGCCCAAGACAAAGTGGGTGAA	1590
QY	1532	CCATGAAGACCTGTGTGTGGGTGAAGAAATTGGAACGGGGGAACTTTGGCGCAAGTTCAG	1591
Db	1591	CCATGAAGACCTGTGTGTGGGTGAAGAAATTGGAACGGGGGAACTTTGGCGCAAGTTCAG	1650
QY	1592	CGGAGCCCTGCGAGCGCAACAACCCCTGGTGGCGGTGAAGTCTTGTGCAAGACGCTCCC	1651
Db	1651	CGGAGCCCTGCGAGCGCAACAACCCCTGGTGGCGGTGAAGTCTTGTGCAAGACGCTCCC	1710
QY	1652	ACCTGACCTCAAGGCCCAAGTTTCTACAGGAAGGAGGATCTTGAAGCATGACCCACC	1711
Db	1711	ACCTGACCTCAAGGCCCAAGTTTCTACAGGAAGGAGGATCTTGAAGCATGACCCACC	1770
QY	1712	CAACATCTGTGCGTCTCATTTGTGTGTGCAACCCAGAAAGACCCCATCTACATCGTCATGA	1771
Db	1771	CAACATCTGTGCGTCTCATTTGTGTGTGCAACCCAGAAAGACCCCATCTACATCGTCATGA	1830
QY	1772	GCTTGTGCAAGGGGGCGAATTCCTCTGACCTTCCCTCCGCAAGGAGGGGGCCGCTGCGGGT	1831
Db	1831	GCTTGTGCAAGGGGGCGAATTCCTCTGACCTTCCCTCCGCAAGGAGGGGGCCGCTGCGGGT	1890
QY	1832	GAAGACTCTGCTGACAGATGTGTGGGGATGCAGCTGTGCGATGAGATACCTGAGAGCAA	1891
Db	1891	GAAGACTCTGCTGACAGATGTGTGGGGATGCAGCTGTGCGATGAGATACCTGAGAGCAA	1950
QY	1892	GTGCTGCATCCACCGGAGACTTGAGTCTCGGAACTGCTCTGTGTGACAGAGAAATGTCT	1951
Db	1951	GTGCTGCATCCACCGGAGACTTGAGTCTCGGAACTGCTCTGTGTGACAGAGAAATGTCT	2010
QY	1952	GAAGATCAAGTGACTTTGGATGTCCCGAAGGAAGGCGGATGGGGTCTATGCAACCTCAGG	2011
Db	2011	GAAGATCAAGTGACTTTGGATGTCCCGAAGGAAGGCGGATGGGGTCTATGCAACCTCAGG	2070
QY	2012	GGGCTCTCAGACAAGTCCCCTGTGAAGTGACCCGACCTGAGGACCTTAACTACGGCCGCTA	2071
Db	2071	GGGCTCTCAGACAAGTCCCCTGTGAAGTGACCCGACCTGAGGACCTTAACTACGGCCGCTA	2130
QY	2072	CTCCTCCGAAGCGAAGTGTGAGACTTTGGCACTTGTCTCTGGGAGACCTTCAAGCCTGGG	2131
Db	2131	CTCCTCCGAAGCGAAGTGTGAGACTTTGGCACTTGTCTCTGGGAGACCTTCAAGCCTGGG	2190
QY	2132	GGCTCTCCCTCATCCCAACTCAGCAATCAGACAGACAGGGAGTTTGTGAGAAAGGGGGG	2191
Db	2191	GGCTCTCCCTCATCCCAACTCAGCAATCAGACAGACAGGGAGTTTGTGAGAAAGGGGGG	2250
QY	2192	CCGTGTGCGCTGCGACAGAGCTGTGTCTGTATGCGCGTTCAGGCTCATGAGAGAGTCTG	2251
Db	2251	CCGTGTGCGCTGCGACAGAGCTGTGTCTGTATGCGCGTTCAGGCTCATGAGAGAGTCTG	2310
QY	2252	GGCTCATGAGCTGTGGACAGCGGCCAGCTTTCAGCAACCATCTACAGAGCTGACAGACAT	2311
Db	2311	GGCTCATGAGCTGTGGACAGCGGCCAGCTTTCAGCAACCATCTACAGAGCTGACAGACAT	2370
QY	2312	CCGAAAGCGGCAATCGGTGAAGGCTGGGACCCCTTCTCAAGCTGTGTGCTCTGACAGGCT	2371
Db	2371	CCGAAAGCGGCAATCGGTGAAGGCTGGGACCCCTTCTCAAGCTGTGTGCTCTGACAGGCT	2430
QY	2372	AGGTGACGTCTCTCAGCGGCTCAGCTCATATGTGACAGCTCTTCAACAGTCCGTGACCTC	2431
Db	2431	AGGTGACGTCTCTCAGCGGCTCAGCTCATATGTGACAGCTCTTCAACAGTCCGTGACCTC	2490
QY	2432	CTGCCACAGCATCCACATGCGCGGAGGATCAGCGCGTGTCTCTGTGTCTCCTGCG	2491

Db 2491 CTGCGACGACATCCACATGCGCGAGATGACGCGCTGCTCTCTGTCCTGC 2550
QY 2492 TGTGTGACGAGGCTTCTCTTCCGAGCAAAACAAATTAACCTTGTGC 2540
Db 2551 TGCCGCAAGGCTTCTCTTCCGAGCAAAACAAATTAACCTTGTGC 2599

RESULT 7
ADH43089
ID ADH43089 standard; cDNA, 2680 BP.
XX ADH43089;
XX
XX
DT 25-MAR-2004 (first entry)
XX
XX CRAM protein related cDNA #SEQ ID 1.
XX
XX Neuroprotective; neutralisation; CRAM;
XX collapsing-response mediator protein-associated molecule; gene therapy;
XX mitochondria; drug development; neural disease; pathosis; gene; ss.
XX Mus musculus.
XX
XX
XX Key Location/Qualifiers
XX CDS 60..2522
XX FT /*tag= a
XX FT
XX
XX PN WO2004001038-A1.
XX PD 31-DEC-2003.
XX PF 19-JUN-2003; 2003WO-JP007766.
XX PR 19-JUN-2002; 2002JP-00179105.
XX
XX (NEW-) NEW IND RES ORG.
XX
XX Yanagi S;
XX
XX WPI; 2004-099123/10.
XX DR P-PSDB; ADH43090.
XX
XX
XX Genes and proteins participating in neutralization of cells or tissues,
XX useful in gene therapy and regeneration medicine, applicable in
XX diagnosis, drug development for neural diseases and study of mechanism of
XX pathosis.
XX
XX Claim 5; SEQ ID NO 1; 101bp; Japanese.
XX
XX
XX The invention relates to a method for inducing neutralisation of cells or
XX tissues by using a protein binding to CRAM (collapsing-response mediator
XX protein-associated molecule) protein or its encoded gene. The proteins
XX and their encoded genes are useful in gene therapy and regenerative
XX CC medicine, e.g. by inducing neutralisation of mitochondria. They are also
XX applicable in diagnosis, drug development for neural diseases and
XX CC studying the mechanism of pathosis. The current sequence represents CRAM
XX protein related cDNA.
XX
XX
XX Sequence 2680 BP; 629 A; 758 C; 791 G; 502 T; 0 U; 0 Other;
XX
XX
XX Query Match 58.0%; Score 1552.2; DB 12; Length 2680;
XX Best Local Similarity 79.1%; Pred. No. 1.6e-200;
XX Matches 1970; Conservative 0; Mismatches 303; Indels 216; Gaps 2;

QY 60 CGGACAGCCTATGAGGCTTCTCTTGAAGCTGACCCCGAGGCCACGGGGTCTTG 119
Db 48 CAGGACACCATTAAGGCTTCTCTTGAAGCTGACCCCGAGGCCACGGGGAGTG 107
QY 120 CAGCAATGAGAGGCGGAGCTTCTCTTGAAGGAGCTGAGAAAGTGAATGCGCCAG 179
Db 108 CAGCAGATGAGAGAGGCTGAGCTGCGCTTATGAGGGGCAATGAGATGCGCCAG 167

QY 180 CGGATCAAGATGACAGGAGATGACAGGACTGCTTCAACACATGCTCCTGACAGACGT 239
Db 168 AGGGTCAAGATGACCGGAAATATGACAGATGCTTCAACACATGCTCCTGACAGACAGC 227
QY 240 GGGGGCCAGAGCGCGGACCATCAGCCCTGACACCCCATATAGTCAATGCTTGGGTGATC 239
Db 228 GAGAGCCAGAGCTGAGAGCGCGGCCCGACACCCCTGTAGCGCAATGCTTGGGAGAGATA 287
QY 300 ACCAGCAAACTGAGGGGCTGAGCGCTTGTGCTGCGGACACGACAGAGATGTAATCTCA 359
Db 288 ACAAGCCAGACCGAGAACTTGAAGCGGGTGTCTGCGGACACATGACAGAAATCTGAATCTCG 347
QY 360 GGGCCCTTGAAGCAAGCTGAGCTGCTCATCCGAGAAACGCGAGACGCTTGGCAAGACTAC 419
Db 348 GGGCCCTTGAAGCAAACTGAGCTGCTGATCCGAGCGGACACAGCTGAGAAAGAGTAC 407
QY 420 AGCGAGCAATGAGCAGAGCTGACAGAGAGCTTCAACAAACCCACAGCCAGAGCATTTGAG 479
Db 408 AACGAGCAATGAGCAGAGCTGACAGAGAGCTTCAACAAACCCACAGAGCATTTGAG 467
QY 480 AAGCTGAAGAGCAGTACCGAGCTTGTGACAGGAGCAGTGCACCAAGCCAAAGCCAGATAC 539
Db 468 AAGCTGAAGACTCAATACCGGACCTTGTATCAGATGACACCCAGGCGCAGGCAATGAC 527
QY 540 CAGAGGCGCCAGCAAAAGACAGGACCGTGAACAGCCAGAGCAAGATATGTGCGAGCTTG 599
Db 528 CAGAAAGCCAGCAAAAGACAGGATCGAGCAAGAGCCAAAGACAAATATGTGCGAGCTTG 587
QY 600 TGGAACTTCTTGTCTACCAACCGCTATGTGTGTGGGCTGTGGGCTTCCGAGCTTACAC 659
Db 588 TGGAACTTCTTGTCTACCAACCGCTATGTGTGTGGGCTGTGGGCTTCCGAGCTTACAC 647
QY 660 CACAGCACCAACCAACAGCTTCTGCTGCGCGGCTGCTGCGGCTCACTGACAGACCTGCAC 719
Db 648 CACACACCAACCAACCGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 707
QY 720 GAGAGATGCTTGTGATCTTGAAGAGATCTGACAGAAATACCTGAGATTAAGACCTTG 779
Db 708 GAGAGATGCTTGTGATCTTGAAGAGATCTGACAGAAATACCTGAGATTAAGACCTTG 767
QY 780 GTGACAGATGAGTGTGTGCTTCAACCGGAGATGTGCTGACGTCTGCGGCAATCCAG 839
Db 768 GTGACAGAGATGAGTGTGCTTCAACCGGAGATGTGCTGACGTCTGCGGCAATCCAG 827
QY 840 CCGTAGGCTGAGTACCAAGGCTTCTGCGACAGTATGAGTCCGACCTGACCTGACCTGACCC 899
Db 828 CCGTAGGCTGAGTACCAAGGCTTCTGCGACAGTATGAGTCCGACCTGACCTGACCTGACCC 887
QY 900 TGTGTACGTTGATGATGCTCACTGCTTGAAGAGGATGAACCGCTGAGGCTTGGGAGCTC 959
Db 888 TGTGTACGTTTGTATGATGCTGCTTCTTGAAGAGGAGGATGAACCGCTGAGGCTTGGGAGCTC 947
QY 960 CAGCTGAACGAGCTGACTGTGAGAGCTGACACACCTGACCTGACCTGACCTGACCTGAG 1019
Db 948 CAGCTGAACGAGCTGACTGTGAGAGCTGACACACCTGACCTGACCTGACCTGAGATGA 1007
QY 1020 CTGAGCTGTGAGCAACGAGATGAGTGTGACAGCGGACAGAGATGATGACAGCTGACACAG 1079
Db 1008 CTGAGCTGTGAGCAACGAGATGAGTGTGACAGCGGACAGAGATGATGACAGCTGACAGCT 1067
QY 1080 GAGCTCCGAGATGAAGAGAGAACACCCACCCCGGAGACGGGTGACGTGCTGAGCAAG 1139
Db 1068 GAGCTCCGAGATGAAGAGAGAACACCCACCCCGGAGACGGGTGACGTGCTGAGCAAG 1127
QY 1140 AGGCAATGCTGACAGAGACCTGACAGAGGCTGACAGTATGCTGTGACAGCCAGGCAAG 1199
Db 1128 AGGCAATGCTGACAGAGACCTGACAGAGGCTGACAGTATGCTGTGACAGCCAGGCAAG 1187
QY 1200 CTGAGGCGCCAGAGAGATGTGCTGACAGCAACGAGCTGAGACACTGTGAGCGGAGAGCC 1259
Db 1188 CTTCAGGCGCCAGAGAGATGTGCTGACAGCAACGAGCTGAGAGAGCTGTGAGCGGAGAGCC 1247
QY 1260 CGGCTGTGCTGCTCTGAGAGATGACCGCACTCCAGCTGTCTTGGAGCAGAGAGCGA 1319


```

Db      1248  CAGCCGTCGCTCTTTCAGAGTGAACCGCATTCACCTCTCCAC-----GGAGCA 1301
Qy      1320  GAGGGGGGAGAGACACCAAGCTGAGATCTTAAAGCCACATCTCAGAAATCTTCCGC 1379
Db      1302  GAGGGGGGAGAGACACCAAGCTGAGATCTTAAAGCCACATCTCAGAAATCTTCCGC 1361
Qy      1380  CCCAAGTTCTC----- 1390
Db      1362  CCCAAGTTCTTATCTCTCCACCACTGACCTTCGCGAGGTGAGAAAGCTCTTAT 1421
Qy      1391  ----- 1390
Db      1422  GAGCAGCTGTGATACCAAGGGGCTATCCCTGGGCAAGGTAGCTGCTTAAGCAC 1481
Qy      1391  ----- 1390
Db      1482  ACTGGGAGCTTCCTGTTGGGAGAGCCAGGCAAGAGATGATGATCTGTGATG 1541
Qy      1391  -----GAACTGTATCCGACTGGAA 1409
Db      1542  TGGGATGSCACACCCGACATTCATCATTCAGTCTTATGATTAACCTGATCGACTGAA 1601
Qy      1410  GGGGAAGGCTTCTCTAGCATTCCTTGTCTCATGACCACTACTGAGACCCAGACCC 1469
Db      1602  GGGGATGGCTTCCCAAGCATCTCTGCTCATCTCACTGCTGCTCTCCAGCACCC 1661
Qy      1470  CTCACCAAGAGAGTGTGTCTGCAAGGGCTGTGCCAAGAGCAAGTGGTGTG 1529
Db      1662  CTTACAAAGAGAGTGTGTGTCTGTTCAAGGGCGGTGCCAAGAGCAAGTGGTGTG 1721
Qy      1530  AACCATGAGAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1589
Db      1722  AAGCAGAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1781
Qy      1590  ACGGAGCGCTGCGAGCCGACCAACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1649
Db      1782  AGTGGCGGCTTCTGTGCAAGCAATACCCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1841
Qy      1650  CGACCTGACCTTCAAGGCGCAAGTTTCTACAGAAAGCAGAGATCTTGAAGCATCAGCCAC 1709
Db      1842  CCACCTGACCTTCAAGGCGCAAGTTTCTGCAAGAAAGCAGAGATCTTGAAGCATCAGCCAC 1901
Qy      1710  CCCAATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1769
Db      1902  CCCAATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1961
Qy      1770  GAGCTTGTGAGAGGGGGGAGCTTCTGACCTTCTCGGACGAGGGGGGCGGCGCGGG 1829
Db      1962  GAGCTTGTGAGAGGGGGGAGCTTCTGACCTTCTCGGACGAGGGGGGCGGCGCGGG 2021
Qy      1830  GTGAAGACTCTGTGTGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1889
Db      2022  GTGAAGACTCTGTGTGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2081
Qy      1890  AAGTGTGATCAACCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1949
Db      2082  AAGTGTGTATCCACAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2141
Qy      1950  CTGAAGATCAGTGAATTTGGAGATGTCCCGAGAGAGAGCGAGTGTATGAGAGCTCA 2009
Db      2142  CTGAAGATCAGGAGCTTTGGAGATGTCCCGAGAGAGAGCTGTATGAGAGCTGTG 2201
Qy      2010  GGGGGGCTCAGACAAAGTCCCGTGAAGTGAACGACCTGAAGGCTCTTAAGTGAAGGCGC 2069
Db      2202  AGCGGCTCAGACAAAGTCCCGTGAAGTGAACGACCTGAAGGCTCTTAAGTGAAGGCGC 2261
Qy      2070  TACTCTCCGAAAGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2129
Db      2262  TACTCTCAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2321
Qy      2130  GGGGCTCTCCCTATCCCAAGCTCAGCAATCAGACAGCGGAGTTGTGTGAAGGGG 2189

```

```

Db      2322  GGGGCTCAGCCCTATCCCAACCTCAACATCAGACAGACAGGAGTGTGTGAAGAAAGGG 2381
Qy      2190  GGGGCTCTGCTCCCTGCCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2249
Db      2382  CATGCTCTGCTTGTGCCAAGAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2441
Qy      2250  TGGGCTTATGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2309
Db      2442  TGGGCTTATGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2501
Qy      2310  ATCCGAAAGCGGCAATCGATGAGGCTGGGA 2338
Db      2502  ATCCGAAAGCGGCAATCGATGAGGCTGGGA 2530

RESULT 8
AAN70060
ID      AAN70060 standard; cDNA; 856 BP.
XX
AC      AAN70060;
XX
XX      27-AUG-2003 (revised)
DT      25-MAR-2003 (revised)
DT      27-JAN-1991 (first entry)
XX
DE      fes/fps proto-oncogene-related DNA.
XX
XX      Furin; fur gene; tumour diagnosis; ss.
XX
OS      Homo sapiens.
XX      Felle catus.
XX
FH      Key Location/Qualifiers
FT      mat_peptide 3..636
FT      /*tag= a
FT      /product= "fes/fps proto-oncogene protein"
XX
PN      EP246709-A.
XX
XX      25-NOV-1987.
PD      XX
XX      19-MAY-1987; 87EP-00200940.
PF      XX
PR      20-MAY-1986; 86NL-00001271.
XX
PA      (UYXA-) UNIV STICHTING KATHOLIEKE.
XX
PI      Vandeven WJM, Roebroek AJM, Schalken JA;
XX      WPI; 1987-328946/47.
DR      P-PSDB; AAP70055.
XX
PT      Recombinant DNA contg. the fur gene - used for producing furin protein
XX      and antibodies and as a diagnostic aid in the detection of tumours.
PS      Disclosure; Fig 3; 24pp; English.
XX
XX      The fur gene, encoding furin, is in the human and cat genomes directly
CC      upstream of this fes/fps proto-oncogene sequence. Furin is strongly
CC      expressed in specific types of tumours and labelled RNA or DNA probes of
CC      the fur gene and antibodies against furin can be used for diagnostic
CC      purposes. (See also AAN70061-62 and AAP70056). (Updated on 25-MAR-2003 to
CC      correct PA field.) (Updated on 25-MAR-2003 to correct PI field.) (Updated
CC      on 27-AUG-2003 to correct OS field.)
XX
SQ      Sequence 856 BP; 164 A; 265 C; 259 G; 168 T; 0 U; 0 Other;
Query Match 32.0%; Score 856; DB 1; Length 856;
Best Local Similarity 100.0%; Pred. No. 1,1e-106;
Matches 856; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1693  TGAGAGATGACGACCAACCCCAATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1752

```

```

Db      1  TGAAGAGTACAGACCAACCAATGTCGTCTATGTCGTCTGACCCAGAAAGACG  60
Qy      1753 GCATCTACATTCGTCTATGAGCTTTGTCAGAGGGGCGAATTCCTGACCTTCCTCCGACGG  1812
Db      61  CCATCTACATTCGTCTATGAGCTTTGTCAGAGGGGCGAATTCCTGACCTTCCTCCGACGG  120
Qy      1813 AGGGGGCCCGCTGCGGGGTGAAGACTCTGCTGAGATGATGAGGGATGAGAGCTGTGCGA  1872
Db      121  AGGGGGCCCGCTGCGGGGTGAAGACTCTGCTGAGATGATGAGGGATGAGAGCTGTGCGA  180
Qy      1873 TGAAGTACCTGAGAGCAAGTGTGATCACCAGGACCTTGCTGCTGCGAACTGCTGCG  1932
Db      181  TGAAGTACCTGAGAGCAAGTGTGATCACCAGGACCTTGCTGCTGCGAACTGCTGCG  240
Qy      1933 TGAAGAGTAAAGATGTCCTGAAGATGATGATGATGATGATGATGATGATGATGATG  1992
Db      241  TGAAGAGTAAAGATGTCCTGAAGATGATGATGATGATGATGATGATGATGATGATG  300
Qy      1993 GGGTCTATGACAGCTTCAGAGGGGCTGACAGCAAGTCCCGTGAAGTGAACCGACCTGAG  2052
Db      301  GGGTCTATGACAGCTTCAGAGGGGCTGACAGCAAGTCCCGTGAAGTGAACCGACCTGAG  360
Qy      2053 CCCTTAACACGCGCCCTACTCTCTCCGAAAGCGAGCTGTGAGCTTTGGCATTTGCTCT  2112
Db      361  CCCTTAACACGCGCCCTACTCTCTCCGAAAGCGAGCTGTGAGCTTTGGCATTTGCTCT  420
Qy      2113 GGGAGACCTTCAAGCTTCAGGGGGCTCCCGCTATCCCAACCTCAAGCATCAAGAGACG  2172
Db      421  GGGAGACCTTCAAGCTTCAGGGGGCTCCCGCTATCCCAACCTCAAGCATCAAGAGACG  480
Qy      2173 AGTTTGTGAGAGAGGGGGGCGCTCTGCTCCCTGACCAAGCTGTGTCTGTGATGCGTGT  2232
Db      481  AGTTTGTGAGAGAGGGGGGCGCTCTGCTCCCTGACCAAGCTGTGTCTGTGATGCGTGT  540
Qy      2233 GGCTCATGAGACAGTCTGAGGCTTATGAGCTTGAGGCGAGGCGGACCTTCAAGCATCT  2292
Db      541  GGCTCATGAGACAGTCTGAGGCTTATGAGCTTGAGGCGAGGCGGACCTTCAAGCATCT  600
Qy      2293 ACCAGAGCTGACAGACATCCGAAGCGGATGAGAGCTGTGAGAGCCCTTCTCAAGC  2352
Db      601  ACCAGAGCTGACAGACATCCGAAGCGGATGAGAGCTGTGAGAGCCCTTCTCAAGC  660
Qy      2353 TGGTGGCTCTGAGAGCTTATGAGCTTCTGAGGCGCTCCAGCTCATATGCTGACAGC  2412
Db      661  TGGTGGCTCTGAGAGCTTATGAGCTTCTGAGGCGCTCCAGCTCATATGCTGACAGC  720
Qy      2413 TCTTCAAGTCTGAGACTCTGTCACACAGATCAACACTGCGGCGAGATGACGCCCT  2472
Db      721  TCTTCAAGTCTGAGACTCTGTCACACAGATCAACACTGCGGCGAGATGACGCCCT  780
Qy      2473 GTCCTCTGTGTCCCTGCTGCTGCGAGGCTTCTTCCGGGCGAAGAACATTAACCC  2532
Db      781  GTCCTCTGTGTCCCTGCTGCTGCGAGGCTTCTTCCGGGCGAAGAACATTAACCC  840
Qy      2533 ACTTGTGCCCACTGA 2548
Db      841  ACTTGTGCCCACTGA 856

```

RESULT 9
ADL71052

ID ADL71052 standard; DNA; 2818 BP.

XX ADL71052;

XX 20-MAY-2004 (first entry)

DE Gene encoding type II collagen expression promoting protein, seq id 45.

KW Osteopathic; antiinflammatory; antiarthritis; antiarthritic;
KW gene therapy; type II collagen; expression; cartilage disease;
KW osteoarthritis; cartilage defect; rheumatoid arthritis; human; gene; ds.

```

OS Homo sapiens.
XX
PN MO2003087375-A1.
XX
PD 23-OCT-2003.
XX
PF 16-APR-2003; 2003WO-JP004802.
XX
PR 16-APR-2002; 2002JP-0013908.
PR 19-APR-2002; 2002US-0373594P.
XX
PA (ASAH ) ASAH KASEI KK.
XX
PI Matsuda A, Honda G, Muramatsu S;
XX
DR MPI: 2003-845331/78.
DR P-PSDB; ADL71053.
XX
PT New purified protein that promotes type II collagen expression, useful
PT for preventing and treating a cartilage disease, e.g. osteoarthritis,
PT cartilage defect, or rheumatoid arthritis.
XX
PS Claim 4; SEQ ID NO 45; 271bp; English.
XX
CC The invention relates to a purified protein (I) that promotes type II
CC collagen expression. Also disclosed is an isolated polynucleotide
CC encoding (I), a recombinant vector comprising the polynucleotide and a
CC gene therapeutic agent comprising the recombinant vector as an active
CC ingredient. The proteins, genes, agents and methods are useful for
CC preventing and treating a cartilage disease, e.g. osteoarthritis,
CC cartilage defect, or rheumatoid arthritis. The current sequence
CC represents a human gene of the invention encoding a protein that promotes
CC type II collagen expression.
XX
SQ Sequence 2818 BP; 909 A; 558 C; 684 G; 667 T; 0 U; 0 Other;
XX
Query Match 14.0%; Score 375; DB 11; Length 2818;
Best Local Similarity 64.8%; Pred. No. 7.4e-42;
Matches 606; Conservative 0; Mismatches 320; Indels 9; Gaps 3;
Qy 1387 TCTCGAATCTGTACCGACCTGGAAGGGAGCTTCTAGCATTCCTTGTCTATGAGAC 1446
Db 1686 TCGATATATCTGTATGATTCGAGGACCGGGTTTCAAAATCCCAAGCTTATGATC 1745
Qy 1447 ACTTACTGAGCACCGGACGCCCTTCAACCAAGAGAGTGTGTGCTGTCAGAGGGCTG 1506
Db 1746 ACCATCTCATATCAANAGCAAGTATCATCAAGAGATGTTTCATTGGGAGAAATTA 1805
Qy 1507 TGCCCAAGGACAAG--TGGGTGCTGAACCATGAGACCTGTGTGGTGAAGCAATTG 1563
Db 1806 TCCCAAGGATTAAGAAATGGTTCTCAATCATGAAGATGTTTCATTGGGAGAAATTA 1865
Qy 1564 GACGGGGGAATTTGGCGAAGTGTTCAGGAGCGCTGGAAGCCGACACACCTGTGTG 1623
Db 1866 GCAAGGGGAATTTGGTGAAGGTATMAAG--GCACATTAAGATTAAGATCCCTGTG 1922
Qy 1624 CGGTGAAGTCTGTGAGAGAGCGCTCCAGCTGACCTCAAGGCGAAGTTTCTACAGAG 1683
Db 1923 CCATTAAACGTGCAAGAGAGACCTTCTCAGGAATTAATAAGTTTCTACAGAG 1982
Qy 1684 CGAGATCTTGAAGAGTACAGCAACCCCAACATGCTGCTCATGTGTCTGACCC 1743
Db 1983 CCAAATCTGAAGCAATATGATCATCCCAATATTTCAAACTGATAGGCGTGTGACAC 2042
Qy 1744 AGAAGAGCCCATCTACATGCTCATGAGAGCTTGTGAGGGGGCGACTTCTGACCTTCC 1803
Db 2043 AAAGCAGCTGTCTACATCATTAATGAGAACTGTCCAGAGGGGTGATTTTCTGACATTC 2102
Qy 1804 TCGCAGCGAGGGGGCGGCTGCGGGTGAAGACTGTGTCAGATGATGAGGGAGTGCAG 1863
Db 2103 TGAAGAGGAGAGAGCGAGCTGAAGCTGAAGAGTGTGTGATTTTCTTGGAGTTG 2162
Qy 1864 CTGCTGTCATGAGTACTGTGAGAGCAAGTGTGATCCACCGGAGCTGGCTGCGA 1923

```

Db	2163	CTGCTGGCATTCTTGTATCTCGAGAGCAAGAACTGCATTCAAGGGACCTGGCTGCACGGA	22222
Qy	1924	ACTGCTGGTGACAGAGAGAATGTCCTGAAGATCACTGACTTTTGGATGTGCCGAGAGG	19833
Db	2223	ACTGCTGGTGAGTGAAGAAATAATCTGAAAATCAAGTACTTTGGAAATGTCCTGGCAAG	22822
Qy	1984	AAGCCGATGGGGTCTATGACAGCTTCAGGGGGCTCAGACAAAGTCCCGTGAATGAGACG	20433
Db	2283	AAGCGGTGGAGTGTATTCATCTTC--TGCTTTAAGCAATTCCTATTAAATGACAG	23339
Qy	2044	CACCTGAGGCCCTTAACTAGAGCCGCTACTCTCCGAAAGGAGAGTGTGAGCTTTGGCA	21033
Db	2340	CACCGAAGCTCTTAATTATTTAGAGATACAGTTCGAAAGTGAAGCTGTGAGCTTCGGCA	23999
Qy	2104	TCTTGCTCTGGGAGACCTTCAGCCTGGGGGGCTCCCCCTATCCCAACTCAGCATCAGC	21633
Db	2400	TCTCTCTCTGGGAACCTTCAGCCTTAGAGTCTGCCCTTACCTCTGGAATGACAAACAGC	24599
Qy	2164	AGACACGGGAGTTTGTGGAGAAGGGGGGCGTCTGCCCTGCCAGAGCTGTGTCTGTATG	22233
Db	2460	AAGCGCGAGACAGTGGAGAGAGGATACGGGAATGTCAAGCCCAAGAACTGTCCAGAGG	25119
Qy	2224	CCGTGTTCAAGGCTCATGAGCAATGCTGTGGGCTTATGAGCCTTGGGCAAGCGGCCAGCTTCA	22833
Db	2520	AGGTTTTTACATCATGATGAAGTGTGGGATTCAAGCCCTGAAAACCGCCTTAAGTTCA	25799
Qy	2284	GCACCATCTACAGAGAGCTGACAGAGCAATCCGAAAAG	2318
Db	2580	ACGACCTTCAACAAAGAGCTCACTGTCAATCAAGAAAG	2614

XX	RESULT 10
XX	AD054298
XX	ID AD054298 standard; DNA; 527 BP.
XX	AC AD054298;
XX	DT 21-OCT-2004 (first entry)
XX	DE Novel canine microarray-related DNA sequence SeqID5600.
XX	KM canine microarray; drug screening; toxicity assay;
XX	KM environmental pollutant; cellular response; gene expression profile;
XX	KM toxic response; liver necrosis; fatty liver disease;
XX	KM protein adduct formation; hepatitis; dog; ds.
XX	OS Canis familiaris.
XX	OS WO2004063324-A2.
XX	PN 29-JUL-2004.
XX	PD 05-MAY-2003; 2003WO-US013853.
XX	PF 03-MAY-2002; 2002US-0377240P.
XX	PR (GENE-) GENE LOGIC INC.
XX	PA (PFIZ) PFIZER PROD INC.
XX	PI DiSgane JC, Porter M, Wei T;
XX	DR WPI; 2004-561890/54.
XX	PT New isolated nucleic acid molecule, useful for drug screening and
XX	PT toxicity assays or for assessing the impact, including toxicity, of a
XX	PT compound, pharmaceutical agent or environmental pollutant on a cell or
XX	PT living organism.
XX	PS Claim 1; SEQ ID NO 5600; 41pp; English.
XX	CC This invention is related to a novel isolated canine nucleic acid
XX	CC sequences and the construction of canine microarrays containing a

significant portion of the canine genome. The isolated canine nucleic acid sequences of the invention may be useful for drug screening and toxicity assays. The invention is therefore useful for assessing the impact, including toxicity, of a compound, pharmaceutical agent or environmental pollutant on a cell or living organism. The methods are useful for detecting genes that are up- or down-regulated in canines in a disease state. The sequences are useful as diagnostic agents or markers to detect a cellular response in a sample individually or as part of a gene expression profile. It is also useful as a target for agents that modulate gene expression or activity. The database is useful for producing electronic Northern blot that allow the user to determine the cell type or tissue in which a given gene is expressed and to allow determination of the abundance or expression level of a given gene in a particular tissue or cell. The methods are useful for determining the similarity of a toxic response to one or more individual compounds. The methods are useful for predicting at least one toxic response or the likelihood that a compound or test agent will induce various specific pathologies such as those of the liver (liver necrosis, fatty liver disease), protein adduct formation or hepatitis), those of the kidney, heart, brain or testes, or other pathologies associated with at least one of the toxins. The methods are also useful for predicting or elucidating the potential cellular pathways influenced, induced or modulated by the compound or test agent due to the similarity of the expression profile compared to the profile induced by a known toxin. The present sequence is that of a canine DNA sequence which was claimed for use during the production of a canine microarray of the invention.

Sequence 527 BP; 97 A; 157 C; 162 G; 98 T; 0 U; 13 Other;

Query Match	14.0%;	Score 374;	DB 13;	Length 527;
-------------	--------	------------	--------	-------------

DESeq local similarity: 0.928; Freq: NO: 10-14,
Matches 406; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

Qy	1215	CTGTGCGAATGCTGTGTGTGACAGAGAAATGTCTGGAATCATGTGACTTTGGGATGT	197
Db	1	CTGCCCCGAACTGCTGTGTGTGACAGAGAAATGCTCTGAAATCATGCACTTTGGCATGT	60
Qy	1975	CCCGAGAGAAAGCCGATGAGGTCTTATGCAAGCTTCAGGGGGCTCAGCAATGCCCTGTGA	203
Db	61	CCCGGAGAAAGACTGATGTGACTATGTGAGGCTTCAGGGGGCTCAGCAATGCTCTGTGA	120
Qy	2035	AGTGGACCGCACTTGAGAGGCTTTAACTAAGGCGGCTACTCCTCCGAAAGGCAAGCTGTGTGA	209
Db	121	AGTGGACAGCAACCNAGGCTTTTAACTAAGGACGCTACTCCTCCGAGGTGACGTGTGA	180
Qy	2095	GCTTTGGCATTTTGCTCTGGGAGACCTTCAGGCTGGGGGCTCCGCCATGCCAACTCA	215
Db	181	GCTTTGGCATTTTGCTCTGGGAGGCTTCAGGCTTAAGAGCTTCCGCCATGCCAACTCA	240
Qy	2155	GCAATCAGCAGACACGGGAGTTTGTGAGAAAGGGGGCGGTCGTGCTCCCTGCCAAGCTGT	221
Db	241	GCAATCAGCAGACCCGGAGTTTGTGAGAAAGGGGGCGGCTCCTCCGCCAAGCTGT	300
Qy	2215	GTCTGTGATGCGTGTTCAGGCTCATGTGAGCATGTGCTGGGCTTATGAGGCTGGGAGCGGG	227
Db	301	GCCCTGACGCTGTGTTCAGGCTCATGTGAGCATGTGCTGGGCTTATGAGGCTGGGAGCGGG	360
Qy	2275	CCAAGCTTCAGCACCATCTTACCCAGGAGGTGCAGACATCCGAAAGCGGCAATCGGTGAGGCT	233
Db	361	CCAAGCTTCNNNNNNNNNTATCCAGNNGCTGCAGACATCCGGAGGCGGCAATCGGTGAGGCT	420
Qy	2335	GGGACCCCCCTTCTCAAGCTGTGTGCTCTCTGCAGGC	2369
Db	421	GGG-CCCCCTTCTCAAGCTGTGAGGCTCTGTGGGC	454

RESULT	11
ABK84514	
ID	ABK84514 standard; cDNA; 12263 BP
XX	
AC	ABK84514;
XX	
DT	14-AUG-2002 (first entry)

XX Human cDNA differentially expressed in granulocytic cells #1085.
 DE
 XX
 KM Human; se; granulocytic cell; DNA chip; bacterial infection;
 KM viral infection; parasitic infection; protozoal infection;
 KM fungal infection; sterile inflammatory disease; psoriasis;
 KM rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KM cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KM adult respiratory distress syndrome; inflammatory bowel disease;
 KM Crohn's disease; ulcerative colitis; periodontal disease;
 KM granulocyte activation; chronic inflammation; allergy.
 XX
 OS Homo sapiens.
 XX
 PN W0200228999-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 03-OCT-2001; 2001MO-US030821.
 XX
 PR 03-OCT-2000; 2000US-0237189P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 XX MPI; 2002-435328/46.
 DR
 XX
 PT Detecting granulocyte activation by detecting differential expression of
 PT gene associated with granulocyte activation, which serves as diagnostic
 PT markers that is useful for monitoring disease states and drug toxicity.
 XX
 PS Claim 1; SEQ ID NO 1085; 1149P; English.
 XX
 CC The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing the
 CC expression level to an expression level in an unactivated GC, where
 CC differential expression of Gs is indicative of GCA. Also included are
 CC modulating (M2) Gs by contacting GC with an agent that alters the
 CC expression of at least one gene in Gs; (2) screening (M3) for an agent
 CC capable of modulating GCA or an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease using the gene expression
 CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease, by detecting the level of
 CC expression in a sample of the tissue of gene(s) from Gs, where the level
 CC of expression of the gene is indicative of inflammation; (4) treating
 CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease, by contacting a tissue having inflammation with an
 CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
 CC is useful for detecting GCA; M2 is useful for modulating Gs; M3 is useful
 CC for screening an agent capable of modulating GCA preferably in an
 CC inflammation in a tissue; M4 is useful for detecting an inflammation
 CC (especially chronic) in a tissue, an allergic response in a subject,
 CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
 CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
 CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
 CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
 CC disease, ulcerative colitis, periodontal disease, also bacterial
 CC infection, viral infection, parasitic infection, protozoal infection,
 CC fungal infection and M5 is useful for treating one of the above
 CC conditions. The present sequence represents a gene differentially
 CC expressed in granulocytes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIP0 at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 12263 BP; 2497 A; 3568 C; 3664 G; 2534 T; 0 U; 0 Other;
 XX
 Query Match 13.6%; Score 362.4; DB 6; Length 12263;

Best Local Similarity 99.7%; Pred. No. 3,6e-40;
 Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2187 GGGGGCCCTTCGCGCCCAAGAGCTGTCTTGATGCGGTTCAGGCTCATGAGCGAG 2246
 DB 11371 GGGGGCCCTTCGCGCCCAAGAGCTGTCTTGATGCGGTTCAGGCTCATGAGCGAG 11430
 QY 2247 TCGTGGGCTTATGAGCTGTGGGAGCGGCCAGCTTCAGACCATTCAGAGAGCTGCAG 2306
 DB 11431 TCGTGGGCTTATGAGCTGTGGGAGCGGCCAGCTTCAGACCATTCAGAGAGCTGCAG 11490
 QY 2307 AGCATTCGAAAGCGGATGCTGAGGCTGGGACCCCTTCTCAAGCTGTGCTTGTGA 2366
 DB 11491 AGCATTCGAAAGCGGATGCTGAGGCTGGGACCCCTTCTCAAGCTGTGCTTGTGA 11550
 QY 2367 GGCCTTAGTGAGCTTCCTTCAGAGGCTCCAGCTCATATGCTGACAGCTTCACAGCTCG 2426
 DB 11551 GGCCTTAGTGAGCTTCCTTCAGAGGCTCCAGCTCATATGCTGACAGCTTCACAGCTCG 11610
 QY 2427 GACTCCTGACACAGATCCAGATCCAGCGGAGATGACAGCGCGCTCTCTGTGTCTC 2486
 DB 11611 GACTCCTGACACAGATCCAGATCCAGCGGAGATGACAGCGCGCTCTCTGTGTCTC 11670
 QY 2487 CCTGCTGTGCGCAGGCGCTTCTCTTCGCGGCAAGAAACAATTAACCACTTGTGCCACTG 2546
 DB 11671 CCTGCTGTGCGCAGGCGCTTCTCTTCGCGGCAAGAAACAATTAACCACTTGTGCCACTG 11730
 QY 2547 AAAA 2550
 DB 11731 AAAA 11734
 RESULT 12
 ID ABA94501 standard; DNA; 15297 BP.
 XX
 AC ABA94501;
 XX
 DT 09-APR-2002 (first entry)
 XX
 XX Human proto-oncogene tyrosine kinase genomic DNA sequence.
 DE
 XX
 KM Proto-oncogene tyrosine kinase; ptk; tumour; cytostatic; anti-leukemic;
 KM gene therapy; protein therapy; vaccine; enzymatic-inhibition; human;
 KM anti-kinase; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 2563..13081
 FT /tag= a
 FT /product= "ptk"
 FT /note= "contains introns"
 FT /tag= b
 FT /number= 1
 FT intron 2776..2927
 FT /tag= c
 FT /number= 1
 FT /number= 1
 FT exon 2928..3101
 FT /tag= d
 FT /number= 2
 FT intron 3102..4476
 FT /tag= e
 FT /number= 2
 FT /number= 2
 FT exon 4477..4573
 FT /tag= f
 FT /number= 3
 FT intron 4574..4702
 FT /tag= g
 FT /number= 3
 FT exon 4703..4886
 FT /tag= h

FT /number= 4
 FT intron 4887..6827
 FT /tag= 1
 FT /number= 1
 FT exon 6828..6965
 FT /tag= 4
 FT /number= 4
 FT /tag= j
 FT /number= 5
 FT intron 6966..7038
 FT /tag= k
 FT /number= 5
 FT exon 7039..7158
 FT /tag= 1
 FT /number= 6
 FT intron 7159..7361
 FT /tag= m
 FT /number= 6
 FT exon 7362..7484
 FT /tag= n
 FT /number= 7
 FT intron 7485..7614
 FT /tag= o
 FT /number= 7
 FT exon 7615..7801
 FT /tag= p
 FT /number= 8
 FT intron 7802..7923
 FT /tag= q
 FT /number= 8
 FT exon 7924..8007
 FT /tag= r
 FT /number= 9
 FT intron 8008..9075
 FT /tag= s
 FT /number= 9
 FT exon 9076..9198
 FT /tag= t
 FT /number= 10
 FT intron 9199..9579
 FT /tag= u
 FT /number= 10
 FT exon 9580..9633
 FT /tag= v
 FT /number= 11
 FT intron 9634..10229
 FT /tag= w
 FT /number= 11
 FT exon 10230..10348
 FT /tag= x
 FT /number= 12
 FT intron 10349..10622
 FT /tag= y
 FT /number= 12
 FT exon 10623..10717
 FT /tag= z
 FT /number= 13
 FT intron 10718..10813
 FT /tag= aa
 FT /number= 13
 FT exon 10814..10937
 FT /tag= ab
 FT /number= 14
 FT intron 10938..11176
 FT /tag= ac
 FT /number= 14
 FT exon 11177..11334
 FT /tag= ad
 FT /number= 15
 FT intron 11335..11458
 FT /tag= ae
 FT /number= 15
 FT exon 11459..11581
 FT /tag= af
 FT /number= 16

FT intron 11582..12938
 FT /tag= ag
 FT /number= 16
 FT exon 12939..13078
 FT /tag= ah
 FT /number= 17
 XX US6340584-B1.
 XX
 XX 22-JAN-2002.
 XX
 XX 27-MAR-2001; 2001US-00817180.
 XX
 XX 27-MAR-2001; 2001US-00817180.
 XX
 XX (PEKE) PE CORP NY.
 XX
 XX Gan W, Ye J, Di Francesco V, Beasley EM;
 XX WPI, 2002-138497/18.
 XX P-PSDB; ABB07354.
 DR
 XX
 XX Nucleic acids encoding a proto-oncogene tyrosine kinase, useful for the
 PT prevention, diagnosis and treatment of e.g. leukemia and lung tumors.
 PT
 XX
 XX Claim 1; Fig 3A-G; 49pp; English.
 PS
 XX
 XX The invention provides isolated nucleic acid sequences encoding a proto-
 CC oncogene tyrosine kinase (pork). The pork polynucleotides and protein may
 CC be used in the prevention, diagnosis and treatment of diseases associated
 CC with inappropriate pork expression, such as lung and kidney tumors,
 CC leukemia and stomach adenocarcinoma. pork may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of pork by expressing
 CC inactive proteins or to supplement the patient's own production of pork.
 CC The encoded pork may be used as an antigen in the production of
 CC antibodies against pork and in assays to identify modulators of pork
 CC expression and activity. The anti-pork antibodies and antagonists may be
 CC used to down regulate expression and activity and as diagnostic agents
 CC for detecting the presence of pork in samples. The present sequence
 CC represents the human pork genomic DNA sequence
 XX
 SQ Sequence 15297 BP; 3267 A; 4306 C; 4478 G; 3246 T; 0 U; 0 Other;
 Query Match 13.6%; Score 362.4; DB 6; Length 15297;
 Best Local Similarity 99.7%; Pred. No. 3.6e-40;
 Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2187 GGGGGCCCTGCTGCCCAAGCTGTGCTGATGCGGTTCAGGCTCATGAGCAG 2246
 DB 12938 GGGGGCCCTGCTGCCCAAGCTGTGCTGATGCGGTTCAGGCTCATGAGCAG 12997
 QY 2247 TGCTGGGCTTATGAGGCTGGGAGCGGCCAGCTTACAGCATTACAGAGCTGAC 2306
 DB 12998 TGCTGGGCTTATGAGGCTGGGAGCGGCCAGCTTACAGCATTACAGAGCTGAC 13057
 QY 2307 AGCATCGAAGGCGCATCGTGAAGCTGGAGCCCTTCTCAAGCTGTGAGCTTGCA 2366
 DB 13058 AGCATCGAAGGCGCATCGTGAAGCTGGAGCCCTTCTCAAGCTGTGAGCTTGCA 13117
 QY 2367 GGCCTAAGTGCAGCTCTTCAAGGCTTCAAGCTTATGTGTGACAGCTTTCAAGTCC 2426
 DB 13118 GGCCTAAGTGCAGCTCTTCAAGGCTTCAAGCTTATGTGTGACAGCTTTCAAGTCC 13177
 QY 2427 GACTCTGCACAGCATCCACATGCGCGGAGATGACAGCCGCTCTCTGTGTC 2486
 DB 13178 GACTCTGCACAGCATCCACATGCGCGGAGATGACAGCCGCTCTCTGTGTC 13237
 QY 2487 CCTGCTGTGCCAGGCTTCTTCTTCCGGGAGAAAACAATTAACATTGTGCCCAC 2546
 DB 13238 CCTGCTGTGCCAGGCTTCTTCTTCCGGGAGAAAACAATTAACATTGTGCCCAC 13297
 QY 2547 AAAA 2550

Db 13298 AACV 13301

RESULT 13
ABV75414
ID ABV75414 standard; DNA; 15297 BP.
XX
XX ABV75414;
XX 24-JAN-2003 (first entry)
XX
XX Human kinase genomic DNA sequence.
XX
XX Human; kinase; proto-oncogene tyrosine kinase; antiarteriosclerotic;
XX cytostatic; antiinflammatory; antiproliferative; gene therapy;
XX protein kinase; drug screening assay; tissue typing; chromosome 15;
XX pharmacogenomic analysis; inflammation; cancer; leukaemia; lung tumours;
XX kidney tumour; stomach adenocarcinoma; arteriosclerosis; psoriasis; SNP;
XX single nucleotide polymorphism; gene; ds.
XX
XX Homo sapiens.
XX
XX Key
XX Location/Qualifiers
XX variation
XX /tag= a
XX /replace(28,T)
XX /standard_name= "Single nucleotide polymorphism"
XX /replace(847,G)
XX /tag= b
XX /standard_name= "Single nucleotide polymorphism"
XX /replace(2159,C)
XX /tag= c
XX /standard_name= "Single nucleotide polymorphism"
XX /replace(2484,T)
XX /tag= d
XX /standard_name= "Single nucleotide polymorphism"
XX /tag= e
XX /product= "protein kinase"
XX /tag= e
XX /number= 1
XX /replace(2577,C)
XX /tag= g
XX /standard_name= "Single nucleotide polymorphism"
XX /tag= h
XX /number= 1
XX /replace(2808,.2809,G)
XX /tag= j
XX /standard_name= "Single nucleotide polymorphism"
XX /replace(2808,T)
XX /tag= i
XX /standard_name= "Single nucleotide polymorphism"
XX /replace(2922,A)
XX /tag= k
XX /standard_name= "Single nucleotide polymorphism"
XX /tag= l
XX /number= 1
XX /replace(4312,T)
XX /tag= n
XX /standard_name= "Single nucleotide polymorphism"
XX /tag= o
XX /number= 3
XX /tag= p
XX /number= 3
XX /tag= q
XX /number= 4
XX /tag= r
XX /number= 4
XX /replace(4903,A)
XX /tag= s
XX /standard_name= "Single nucleotide polymorphism"
XX /replace(5193,G)
XX /tag= t
XX /standard_name= "Single nucleotide polymorphism"
XX /tag= u
XX /number= 5
XX /tag= v
XX /number= 5
XX /tag= w
XX /number= 6
XX /tag= x
XX /number= 6
XX /tag= y
XX /number= 7
XX /tag= z
XX /number= 7
XX /tag= aa
XX /number= 8
XX /tag= ab
XX /number= 8
XX /tag= ac
XX /number= 9
XX /tag= ad
XX /number= 9
XX /tag= ae
XX /number= 10
XX /tag= af
XX /number= 10
XX /tag= ag
XX /number= 11
XX /tag= ah
XX /number= 11
XX /tag= ai
XX /number= 12
XX /tag= aj
XX /number= 12
XX /tag= ak
XX /number= 13
XX /tag= al
XX /number= 13
XX /tag= am
XX /number= 14
XX /tag= an
XX /number= 14
XX /tag= ao

FT /tag= q
FT /number= 4
FT 4887..6827
FT /tag= r
FT /number= 4
FT /replace(4903,A)
FT /tag= s
FT /standard_name= "Single nucleotide polymorphism"
FT /replace(5193,G)
FT /tag= t
FT /standard_name= "Single nucleotide polymorphism"
FT /tag= u
FT /number= 5
FT /tag= v
FT /number= 5
FT /tag= w
FT /number= 6
FT /tag= x
FT /number= 6
FT /tag= y
FT /number= 7
FT /tag= z
FT /number= 7
FT /tag= aa
FT /number= 8
FT /tag= ab
FT /number= 8
FT /tag= ac
FT /number= 9
FT /tag= ad
FT /number= 9
FT /tag= ae
FT /number= 10
FT /tag= af
FT /number= 10
FT /tag= ag
FT /number= 11
FT /tag= ah
FT /number= 11
FT /tag= ai
FT /number= 12
FT /tag= aj
FT /number= 12
FT /tag= ak
FT /number= 13
FT /tag= al
FT /number= 13
FT /tag= am
FT /number= 14
FT /tag= an
FT /number= 14
FT /tag= ao

FT	/number= 15
FT	.1135.
FT	intron
FT	/*tag= ap
FT	/number= 15
FT	.11459.
FT	exon
FT	/*tag= ag
FT	/number= 16
FT	.11582.
FT	intron
FT	/*tag= at
FT	/number= 16
FT	exon
FT	.12939.
FT	/*tag= as
FT	/number= 17
PN	WO200277191-A2.
XX	
PD	03-OCT-2002.
XX	
PF	27-MAR-2002; 2002WO-USO09325.
XX	
PR	27-MAR-2001; 2001US-00817180.
PR	06-DEC-2001; 2001US-00003295.
PA	(PEKE) PE CORP NY.
XX	
PI	Gan W, Ye J, Di Francesco V, Beasley EM;
XX	
DR	WPI; 2003-029927/02.
XX	
PT	P-PSDB; ABB99046.
PT	
PT	New human kinase protein, useful for treating or diagnosing disorders associated with an absence of, inappropriate, or unwanted expression of the protein, e.g., inflammation or cancer, in drug screening assays and pharmacogenomics.
XX	
XX	Claim 4 (b); Fig 3; 75pp; English.
CC	The invention relates to a newly isolated peptide sequence of a human kinase that is related to the proto-oncogene tyrosine Kinase subfamily.
CC	The activity of the kinase of the invention may be described as,
CC	Cytosarctic, antarteriosclerotic, antiinflammatory and antipsoriatic.
CC	Peptides of the invention are useful in assays to determine the biological activity of the protein, in drug screening assays, tissue typing and pharmacogenomic analysis. They are also useful in treating or diagnosing disorders characterised by an absence of, inappropriate, or unwanted expression of the protein, such as inflammation, cancer (e.g., leukaemia, lung tumours, kidney tumours or stomach adenocarcinoma), arteriosclerosis, and psoriasis. Nucleic acid molecules of the invention are useful as probes, primers and chemical intermediates in biological assays. The peptide and nucleic acid sequences are useful as models for the development of human therapeutic targets, aid in the identification of therapeutic proteins and serve as targets for the development of human therapeutic agents that modulate kinase protein activity in cells and tissues that express the protein. The protein of the invention may also be useful in gene therapy. The gene encoding the protein of the invention has been localised to human chromosome 15. The current sequence represents the human Kinase genomic DNA sequence
SQ	Sequence 15297 BP; 3267 A; 4306 C; 4478 G; 3246 T; 0 U; 0 Other;
Query Match	13.6%; Score 362.4; DB 8; Length 15297;
Best Local Similarity	99.7%; Pred. No. 3.6e-40;
Matches 363; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Dn	12938 GGGGCGGTGCGCCTGCCAGACCTGTGTCATATCCCGTTCAGGCTCATGAGACAG 1299
Oy	2187 GGGGCGGTGCGCCTGCCAGACCTGTGTCATATCCCGTTCAGGCTCATGAGACAG 2246
Dn	12938 GGGGCGGTGCGCCTGCCAGACCTGTGTCATATCCCGTTCAGGCTCATGAGACAG 1299
Oy	2247 TGCTGGGCTTGAGACCTGAGGACGGGCCAGGCTTCAGACACATCTACAGAGACTGCAG 2306
Dn	12998 TGCTGGGCTTGAGACCTGAGGACGGGCCAGGCTTCAGACACATCTACAGAGACTGCAG 1305
Oy	2307 AGCATCCGAAGCGGATCGGTGAGGCTGGAGCCCTTTCAAGCTGTGGGCTTGCAG 2366

ID	Accession	Gene	Location/Qualifiers	Position
D6	13058	AGCATCCGAAAGCGGCATCGTGAAGCTGGACCCCTTTCTCAAGTGtggccttcgca		13117
OY	2367	Ggccttagtgacagctctctccagcggtctccagctcatatgctgaagcttttcacagtcttg		2426
D6	13118	Ggccttagtgacagctctctccagcggtctccagctcatatgctgaagcttttcacagtcttg		13177
OY	2427	GACTCCTGCCACACAGATTCACACTCCCGGAGGATGACGCGCGGTCTCTCTGTGTGTC		2486
D6	13178	GACTCCTGCCACACAGATTCACACTCCCGGAGGATGACGCGCGGTCTCTCTGTGTGTC		1323
OY	2487	CCTGCTGCTCCAGAGGCTTCTCTTCCTCCGGGACAGAAACAATAAACCATTGTGCCCACCTG		2546
D6	13238	CCTGCTGCTCCAGAGGCTTCTCTTCCTCCGGGACAGAAACAATAAACCATTGTGCCCACCTG		1329
OY	2547	AAAA 2550		
D6	13298	AACA 13301		
RESULT 14				
ID	ADM28582	standard; cDNA; 15297 BP.		
XX	ADM28582;			
XX	ADMD8582;			
XX	17-JUN-2004	(first entry)		
DE	Human proto-oncogene tyrosine kinase DNA.			
XX	des; gene therapy; human; proto-oncogene; tyrosine kinase; cancer.			
XX	Homo sapiens.			
OS				
XX				
FH	Key	Location/Qualifiers		
FT	variation	/+tag= a replace(28,t)		
FT		/note= "Single nucleotide polymorphism"		
FT	variation	/+tag= b replace(847,g)		
FT		/note= "Single nucleotide polymorphism"		
FT	variation	/+tag= c replace(2159,c)		
FT		/note= "Single nucleotide polymorphism"		
FT	variation	/+tag= d replace(2484,t)		
FT		/note= "Single nucleotide polymorphism"		
FT	CDS	2563..13081		
FT		/+tag= e		
FT	exon	2563..2775		
FT		/+tag= f		
FT	variation	/+tag= g replace(2577,c)		
FT		/note= "Single nucleotide polymorphism"		
FT	intron	2776..2927		
FT		/+tag= h		
FT	variation	/+tag= i replace(2807..2809,gs)		
FT		/note= "Single nucleotide polymorphism"		
FT	variation	/+tag= j replace(2808,t)		
FT		/note= "Single nucleotide polymorphism"		
FT	variation	/+tag= k replace(2922,a)		
FT		/note= "Single nucleotide polymorphism"		
FT	exon	2928..3101		
FT		/+tag= l		
FT	intron	3102..4476		
FT		/+tag= m		
FT	variation	/+tag= n replace(4312,t)		
FT		/note= "Single nucleotide polymorphism"		
FT	exon	4477..4573		

PF 16-APR-2003; 2003WO-JP004802.
XX 16-APR-2002; 2002JP-00113908.
PR 19-APR-2002; 2002US-0373594P.
XX (ASAH) ASAH KASEI KK.
XX Matsuda A, Honda G, Muramatsu S;
PI MPI; 2003-045331/78.
DR P-PSDB; ADL71055.
XX
XX New purified protein that promotes type II collagen expression, useful
PT for preventing and treating a cartilage disease, e.g. osteoarthritis,
XX cartilage defect, or rheumatoid arthritis.
XX
XX Claim 4; SEQ ID NO 47; 271bp; English.
XX
XX The invention relates to a purified protein (I) that promotes type II
CC collagen expression. Also disclosed is an isolated polynucleotide
CC encoding (I), a recombinant vector comprising the polynucleotide and a
CC gene therapeutic agent comprising the recombinant vector as an active
CC ingredient. The proteins, genes, agents and methods are useful for
CC preventing and treating a cartilage disease, e.g. osteoarthritis,
CC cartilage defect, or rheumatoid arthritis. The current sequence
CC represents a human gene of the invention encoding a protein that promotes
CC type II collagen expression.
XX
SQ Sequence 2584 BP; 922 A; 446 C; 565 G; 651 T; 0 U; 0 Other;

Query Match 12.3%; Score 328.8; DB 11; Length 2584;
Best Local Similarity 62.0%; Pred. No. 1.3e-35;
Matches 573; Conservative 0; Mismatches 342; Indels 9; Gaps 3;

QY 1392 AACCTGACCGATGGAAGGAGGCTTCTAGCATCTTGTGCTCATGACCACTA 1451
DB 1552 AACATGATGATTCGAGGCGACTGGGTTTCAAACTTCTCATCTTAATGATCATC 1611
QY 1452 CTGAGCACCAGACCCCTCACCAGAGAGTGTGTCTCTGACAGAGGCTGTGCC 1511
DB 1612 TATACAAACAAACAGGTCATCAAGAAATCAGTGTAGTCTGCTGAATCCTATTC 1671
QY 1512 AAGGACAG---TGGGCTGGAACACATGAGACCTGTGTGGTGAGCAGATTGACGG 1568
DB 1672 AAGGACAGAAATGATTCAGTCATGAAAGATCATATTGAGAGAAATTAATTA 1731
QY 1569 GGGACCTTGGCGAAGTGTCAAGCGAGCGCTCGAGCCGACCAACCTGTGGCGGTG 1628
DB 1732 GGAATTTTGGTGAAGTATTAAG---GCACATTAAGATTAATACTTGTCTGCTG 1788
QY 1629 AAGCTTTTGGAGAGAGCTTCCCACTGACCTCAAGGCCAAGTTTCTACAGAA 1688
DB 1789 AAAACATGTAAGAAAGATCTTCTCAGGAATGAAATTAATAATTTTACAAAG 1848
QY 1689 ATCTGAAGAGTACAGCCACCCCAATGTGTGCTCATTTGTGTCTGACCCAGAG 1748
DB 1849 ATTCTCAAGCAATATGATCATCCCAATATGTCAAACTTAATGAGATTGACACA 1908
QY 1749 CAGCCCATCTACATCTGTCATGAGCTTGTGAGGGGGGCGACTTCTGACCTTCC 1808
DB 1909 CAGCCCTGTCTACATCTTATGAACTGTGTCAGAGGTGATTTCTCACCCTTCT 1968
QY 1809 ACGGAGGGGGCCCGCTGCGGGTGAAGACTGTGTCAGATGTGGGGATGCACTG 1868
DB 1969 AGGAGAGAGATGAATACTAAACATCAACAGTATGAAATTTTCATTAGACGCT 2028
QY 1869 GGCATGAGTACTGAGAGACAAAGTGTGATCCACCGGACCTGTGCTCGGAACTGC 1928
DB 2029 GGTATGTTTATCTCGAGAGTAAAACTGTATACAGGGGACCTTGTGCAAGAA 2088
QY 1929 CTGGTGACAGAGAAATGTCTGTAAGATCAGTGACTTTGGATGTCCGAGAGAAC 1988
DB 2089 CTGGTAGGTGAATAATATGTTCTGAAATCAGTGAATTTGGAATGTCTCTCAAG 2148

QY 1989 GATGGGCTATATGACACCTCAGAGGGGCTCAGACAAAGTCCCGTGAATGAGACCGACCT 2048
DB 2149 GGTGAGTGTATTCATCTTC--TGCTTAAAGCAATTCCTCAATTAATGAGACGACCG 2205
QY 2049 GAGGCCCTTAACATGAGCGGCTACTCTCCGAAAGGAGAGTGTGAGCTTTGGCATCTTG 2108
DB 2206 GAGGCTTTAATTTATGAGAGATACAGTTTCAAGAGAGTGTGAGCTTTGGCATCTT 2265
QY 2109 CTCTGGAGACCTTCAGCCTGGGGGCTCCCTATCCAACTCAGCAATCAGACAGACA 2168
DB 2266 CTCTGGAGACCTTCAGCCTTAGGGGTTTGTCCGTACCTGGAATGACAAATCAGACAGA 2225
QY 2169 CGGAGTTTGTGAGAGAGGGGGGCGCTGTCCCTGCCAGAGCTGTGTCTGATGCCGTG 2228
DB 2326 AGAGACCAAGTAAAGAGATACCGAATGTCAAGTCTCCAGACATGTCCAGAGATAT 2385
QY 2229 TTGAGCTATGAGAGAGTGTGGGCTATGAGCTTGGGAGGGGCCAGCTTCAGCACCC 2288
DB 2386 TCCAAATATCATGATGAAAGTGTGGGATTATTAACCTGAAATATGCCCTTAAGTTCA 2445
QY 2289 ATCTACAGAGCTGACAGACATC 2312
DB 2446 CTTCAGAAAGAGCTCACTATCATC 2469

Search completed: March 19, 2005, 04:31:15
Job time: 1290.71 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 19, 2005, 03:18:16 ; Search time 1122.3 Seconds
(without alignments)
11649.479 Million cell updates/sec

Title: US-10-660-763-1

Perfect score: 2674
Sequence: 1 tccggggtccgcacgcggcc.....aaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenBank: 1: gb_ha: 2: gb_hg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_st: 12: gb_sv: 13: gb_un: 14: gb_vl:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2674	100.0	2674	6	AR183261 Sequence
2	2674	100.0	2674	6	AR456323 Sequence
3	2356	88.1	2788	6	BC035357 Homo sapi
4	2328	87.1	2771	9	HSCFES
5	2326.4	87.0	2889	6	AX281847 Sequence
6	2267	84.8	2697	6	C0725220 Sequence
7	1552.2	58.0	2680	10	MPESCR
8	1024.8	38.3	2715	14	ACFTS140A
9	1024.8	38.3	4788	14	ACF
10	1024.8	38.3	4788	14	ACF
11	856	32.0	856	6	A06935
12	854.4	32.0	4901	13	M22820
13	842.2	31.5	2397	14	FCSTONC
14	831.2	31.1	2946	5	FCGSAONC
15	659.6	24.7	4788	14	BC073445
16	630.8	23.6	2284	14	ACSGAGPS
17	455.2	17.0	772	10	BC038130
18	376.6	14.1	1563	10	BC051249
19	376.6	14.1	2948	10	BC058100

20	376.6	14.1	2994	10	MMU76762	U76762 Mus musculu
21	375	14.0	2069	10	MUSPERT	M32054 Mouse tyros
22	374.8	14.0	1171	10	RNFLK	X13412 Rat mRNA fo
23	362.4	13.6	12263	9	HSPFSPS	X06292 Human c-fes
24	362.4	13.6	15297	6	AR183262	AR183262 Sequence
25	362.4	13.6	15297	6	AR456324	AR456324 Sequence
26	362.4	13.6	33550	9	AC003004	AC003004 Human Chr
27	362.4	13.6	142201	9	AC124248	AC124248 Homo sapi
28	362.4	13.6	156889	2	AC004586	AC004586 Homo sapi
29	342.8	12.8	2947	4	AF187884	AF187884 Canis fam
30	328.8	12.3	2950	9	HUMTKPER	Y03358 Human tyros
31	318.4	11.9	3882	3	SRFESTK	Y17051 Sycon rapha
32	296.2	11.1	361	6	AR175607	AR175607 Sequence
33	296.2	11.1	361	6	AR236724	AR236724 Sequence
34	274.4	10.3	2786	6	AB006567	AB006567 Ephydactia
35	271.6	10.2	3198	6	CQ597087	CQ597087 Sequence
36	271.6	10.2	3369	6	CQ591168	CQ591168 Sequence
37	271.6	10.2	4972	3	BT003462	BT003462 Drosophila
38	268.4	10.0	2303	3	DMU50450	U50450 Drosophila
39	268.4	10.0	3287	3	DMPS85D	X52844 Drosophila
40	230.6	8.6	2537	6	CQ726604	CQ726604 Sequence
41	221.8	8.3	4160	6	A06936	A06936 H. sapiens f
42	221.8	8.3	4160	6	E01406	E01406 DNA sequenc
43	201.2	7.5	2398	9	HSRTKBP	227409 H. sapiens m
44	198	7.4	2955	6	AX481387	AX481387 Sequence
45	198	7.4	3370	6	CQ776660	CQ776660 Sequence

ALIGNMENTS

RESULT 1	AR183261	2674 bp	DNA	linear	PAT 20-APR-2002
LOCUS	AR183261	Sequence 1 from patent US 6340584.			
DEFINITION	Sequence 1 from patent US 6340584.				
ACCESSION	AR183261				
VERSION	AR183261.1	GI:20226854			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 2674)				
AUTHORS	Gan, W., Ye, J., Di Francesco, V. and Beasley, E. M.				
TITLE	Isolated human kinase proteins, and uses thereof				
JOURNAL	Patent: US 6340584-A 1 22-JAN-2002;				
FEATURES	Location/Qualifiers				
source	1..2674				
ORIGIN	/organism="unknown"				
	/mol_type="unassigned DNA"				

Query Match	100.0%	Score 2674;	DB 6;	Length 2674;
Best local Similarity	100.0%	Pred. No. 0;		
Matches 2674;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TCGGGGGTCGCA	CCGGGCTGAGTCGTCGAGCCGCTCCAGAGACAGCTGCCGCTGC	60
DB	1	TCGGGGGTCGCA	CCGGGCTGAGTCGTCGAGCCGCTCCAGAGACAGCTGCCGCTGC	60
QY	61	GGAACGACATAG	GGCTTCTTCTGAGCTGAGCCGCGGCGCAGGGGCTGCTGC	120
DB	61	GGAACGACATAG	GGCTTCTTCTGAGCTGAGCCGCGGCGCAGGGGCTGCTGC	120
QY	121	AGCAATGACAG	AGCCGCTTCTGCTACTGAGAGGACATGAGAAAGTGATGCCAGC	180
DB	121	AGCAATGACAG	AGCCGCTTCTGCTACTGAGAGGACATGAGAAAGTGATGCCAGC	180
QY	181	GGGTCAAGATG	ACAGGAGTATGACAGACTGCTTACCAACATGCTCCCTGACAGACATG	240
DB	181	GGGTCAAGATG	ACAGGAGTATGACAGACTGCTTACCAACATGCTCCCTGACAGACATG	240
QY	241	GGGGCCAGAGC	CGGGCCATGAGCCCTGACAGAGCCCATCACTCACTGCTGAGATCA	300

Db 241 GGGGCCAGAGCCGGGCATCAGCCCTGACGCCCATCAGTCACTGGGCTGAGATCA 300
Qy 301 CCAGCCAAATGAGGGCTGAGCCGCTTGTCTGGGAGACACGAGAGATCTGAATCAG 360
Db 301 CCAGCCAAATGAGGGCTGAGCCGCTTGTCTGGGAGACACGAGAGATCTGAATCAG 360
Qy 361 GGGCCCTGAGCAGCTGAGCTGCTCATCCGGGAAAGGAGCAGCTTCCGAAGACTTACA 420
Db 361 GGGCCCTGAGCAGCTGAGCTGCTCATCCGGGAAAGGAGCAGCTTCCGAAGACTTACA 420
Qy 421 GCGAGCAGTGGCAGCAGCTGACAGAGAGCTCAACAAACCAACAGCAGCATTGAGA 480
Db 421 GCGAGCAGTGGCAGCAGCTGACAGAGAGCTCAACAAACCAACAGCAGCATTGAGA 480
Qy 481 AGCTGAAGAGCCGATACCGAGCTCTGGCACGGGACAGTCCCAAGCCCAAGTAATCC 540
Db 481 AGCTGAAGAGCCGATACCGAGCTCTGGCACGGGACAGTCCCAAGCCCAAGTAATCC 540
Qy 541 AGGAGGCGAGCAAGACCAAGGACCGTGAACAAAGGCAAGATATGTCCGAGCTGT 600
Db 541 AGGAGGCGAGCAAGACCAAGGACCGTGAACAAAGGCAAGATATGTCCGAGCTGT 600
Qy 601 GGAAGCTTTTGTCTCAACACACCGCTATGTGTGGCGGTGCGGAGCTGACACTACCC 660
Db 601 GGAAGCTTTTGTCTCAACACACCGCTATGTGTGGCGGTGCGGAGCTGACACTACCC 660
Qy 661 ACCAGACACCAACCAAGCTCTGCTGCTGCGGCTGCTGCGGTCACTGACAGACTGACG 720
Db 661 ACCAGACACCAACCAAGCTCTGCTGCTGCGGCTGCTGCGGTCACTGACAGACTGACG 720
Qy 721 AGGAGATGAGCTTGCATCTTGAAGAGAGATCTTGAAGAGATCTTGAAGAGATCTG 780
Db 721 AGGAGATGAGCTTGCATCTTGAAGAGAGATCTTGAAGAGATCTTGAAGAGATCTG 780
Qy 781 TGCAGATGAGGTGTGTGGCCATTCACCGGAGATGTGTGAGCTGTCCCGCATCAGC 840
Db 781 TGCAGATGAGGTGTGTGGCCATTCACCGGAGATGTGTGAGCTGTCCCGCATCAGC 840
Qy 841 CTGAGGCTGATGATCAAGAGCTTCTCTGAGACGATGTGGTCCGACCTGAGCTCCACCT 900
Db 841 CTGAGGCTGATGATCAAGAGCTTCTCTGAGACGATGTGGTCCGACCTGAGCTCCACCT 900
Qy 901 GTGTCAAGTTCAGATGATCACTGCTTGAAGAGAGGTGAACCGCTGAGAGCTGGGAGCTCC 960
Db 901 GTGTCAAGTTCAGATGATCACTGCTTGAAGAGAGGTGAACCGCTGAGAGCTGGGAGCTCC 960
Qy 961 AGCTGAACGAGCTGATCTGTGAGAGAGCTGACAGCAACGCTGACCTGACATGAGAGC 1020
Db 961 AGCTGAACGAGCTGATCTGTGAGAGAGCTGACAGCAACGCTGACCTGACATGAGAGC 1020
Qy 1021 TGGCTGTGGCCACCGAGATGTGTTCAGGGGAGAGATGTGTTAGCAGCTTCAACAGG 1080
Db 1021 TGGCTGTGGCCACCGAGATGTGTTCAGGGGAGAGATGTGTTAGCAGCTTCAACAGG 1080
Qy 1081 AGCTCCGAATGAAGAGAGAACCCCAACCCCGGAGAGAGAGCTGCTGGCAAGA 1140
Db 1081 AGCTCCGAATGAAGAGAGAACCCCAACCCCGGAGAGAGAGCTGCTGGCAAGA 1140
Qy 1141 GGGAGTGTCTGCAAGAGACCTGACAGAGGCTGACAGGTAGGCTGTGACGCAAGC 1200
Db 1141 GGGAGTGTCTGCAAGAGACCTGACAGAGGCTGACAGGTAGGCTGTGACGCAAGC 1200
Qy 1201 TGCAGGCGCCAGAGAGATGTCTGACAGCAACGCTGAGAGCACTGGGCGCCGAGAGCC 1260
Db 1201 TGCAGGCGCCAGAGAGATGTCTGACAGCAACGCTGAGAGCACTGGGCGCCGAGAGCC 1260
Qy 1261 CGCTGTGTCTGCTCTGAGAGATGACCGCACTCAGCTGTCTCTGAGAGAGAGAGAG 1320
Db 1261 CGCTGTGTCTGCTCTGAGAGATGACCGCACTCAGCTGTCTCTGAGAGAGAGAGAG 1320
Qy 1321 AGGGGGAGAGAGACCCACGCTGAGATCTTAAAGAGCAATCTCAGAGATCTTCGCC 1380
Db 1321 AGGGGGAGAGAGACCCACGCTGAGATCTTAAAGAGCAATCTCAGAGATCTTCGCC 1380

Db 1321 AGGGGGAGAGAGACCCACGCTGAGATCTTAAAGAGCAATCTCAGAGATCTTCGCC 1380
Qy 1381 CCAAGTTCTCGAACCCTGTAACCGAATGGAAGGGAGGCTTCTAGCATCTCTTGCTCA 1440
Db 1381 CCAAGTTCTCGAACCCTGTAACCGAATGGAAGGGAGGCTTCTAGCATCTCTTGCTCA 1440
Qy 1441 TGCAGCACTTACTGAGACACCCAGAGCCCTCAACAGAGAGAGTGTGTCTGACACA 1500
Db 1441 TGCAGCACTTACTGAGACACCCAGAGCCCTCAACAGAGAGAGTGTGTCTGACACA 1500
Qy 1501 GGGCTGTGCGCAAGAGCAAGTGGTGTGACACATGAGAGCCTGTGTGGTGGAGACA 1560
Db 1501 GGGCTGTGCGCAAGAGCAAGTGGTGTGACACATGAGAGCCTGTGTGGTGGAGACA 1560
Qy 1561 TTGAGAGGGGGAACCTTGGCGAAGTGTTCAGGGAGAGCCTGGAGGCGCAACACCTGG 1620
Db 1561 TTGAGAGGGGGAACCTTGGCGAAGTGTTCAGGGAGAGCCTGGAGGCGCAACACCTGG 1620
Qy 1621 TGGCGGTGAAGTCTTGTGAGAGAGAGCTCCACCTGACCTCAAGGCAAGTTTCTACAG 1680
Db 1621 TGGCGGTGAAGTCTTGTGAGAGAGAGCTCCACCTGACCTCAAGGCAAGTTTCTACAG 1680
Qy 1681 AAGCAGAGATCTTGAAGCAGTACAGCACCACCAACATGCTGTCTCATTTGTGTGCA 1740
Db 1681 AAGCAGAGATCTTGAAGCAGTACAGCACCACCAACATGCTGTCTCATTTGTGTGCA 1740
Qy 1741 CCAGAGAGAGCCCATCTATCATGTGATGAGCTTGTGAGGGGGGCGACTTCTGACCT 1800
Db 1741 CCAGAGAGAGCCCATCTATCATGTGATGAGCTTGTGAGGGGGGCGACTTCTGACCT 1800
Qy 1801 TCTCTCGACAGAGGGGGGCGCTGCGGAGAGACCTGTGTGACAGATGTGTGGAGATG 1860
Db 1801 TCTCTCGACAGAGGGGGGCGCTGCGGAGAGACCTGTGTGACAGATGTGTGGAGATG 1860
Qy 1861 CAGCTGTGCGATGAGATCTCTGAGAGAGATGTCTGATCCACCGGAGACTGTGCTCTC 1920
Db 1861 CAGCTGTGCGATGAGATCTCTGAGAGAGATGTCTGATCCACCGGAGACTGTGCTCTC 1920
Qy 1921 GGAACGTGCTGTGACAGAGAGAAATGCTCTGAAGATCAGTGACTTTGGAGATGCCGAG 1980
Db 1921 GGAACGTGCTGTGACAGAGAGAAATGCTCTGAAGATCAGTGACTTTGGAGATGCCGAG 1980
Qy 1981 AGAAGCCGATGGGCTATATGAGCCTCAAGGGGCTCAGACAAAGTCCCGTGAAGTGA 2040
Db 1981 AGAAGCCGATGGGCTATATGAGCCTCAAGGGGCTCAGACAAAGTCCCGTGAAGTGA 2040
Qy 2041 CCGCACCTGAGGCTTAACTACGGCCGCTACTCTCCGAAAGCGAGCTGTGGAGCTTTG 2100
Db 2041 CCGCACCTGAGGCTTAACTACGGCCGCTACTCTCCGAAAGCGAGCTGTGTGGAGCTTTG 2100
Qy 2101 GGAATCTGTCTGGGAGACCTTCAAGCTGGGGGCTCCGCTATCCCAACTTCAGCAATC 2160
Db 2101 GGAATCTGTCTGGGAGACCTTCAAGCTGGGGGCTCCGCTATCCCAACTTCAGCAATC 2160
Qy 2161 AGCAGACAGGAGATTTGTGAGAGAGGGGGGCGTCTGCTGCTCCAGAGAGTGTCTCG 2220
Db 2161 AGCAGACAGGAGATTTGTGAGAGAGGGGGGCGTCTGCTGCTCCAGAGAGTGTCTCG 2220
Qy 2221 ATGCGGTGTTCAAGCTCATGAGACAGTGTGGGCTATAGAGCTGGGAGCGGCTCAAGT 2280
Db 2221 ATGCGGTGTTCAAGCTCATGAGACAGTGTGGGCTATAGAGCTGGGAGCGGCTCAAGT 2280
Qy 2281 TGCAGACATCTACAGAGAGCTGACAGAGATCGAAGAGGAGCATCGGTGAGGCTGGAGCC 2340
Db 2281 TGCAGACATCTACAGAGAGCTGACAGAGATCGAAGAGGAGCATCGGTGAGGCTGGAGCC 2340
Qy 2341 CCCTTCTCAAGCTGTGTGCTCTGAGGCTTATGATGACACTCTCAGCGGCTCAGACTCA 2400
Db 2341 CCCTTCTCAAGCTGTGTGCTCTGAGGCTTATGATGACACTCTCAGCGGCTCAGACTCA 2400
Qy 2401 TATGTGACAGCTCTTCAAGTCTGAGACTCTGTGACACAGCATTCACACTGCGGAGAG 2460
Db 2401 TATGTGACAGCTCTTCAAGTCTGAGACTCTGTGACACAGCATTCACACTGCGGAGAG 2460

Qy	2461	TTGCAGCGCGCGTGTCTCTCTGTGTGTCCTGTGCTGCTGCAGAGGCTTCTCTCCGGGACGA	2522
Db	2461	ATGCAGCGCCCTGTGTCTCTCTGTGTGTCCTGTGCTGCTGCAGAGGCTTCTCTCCGGGACGA	2520
Qy	2521	AACATATAAACCACTTGTGTGCCCACTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2580
Db	2521	AACATATAAACCACTTGTGTGCCCACTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2580
Qy	2581	AA	2640
Db	2581	AA	2640
Qy	2641	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2674
Db	2641	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2674
RESULT 2			
AR456323		2674 bp	DNA
LOCUS	AR456323		linear
DEFINITION	Sequence 1 from patent US 6686187.		PAT 20-FEB-2004
ACCESSION	AR456323		
VERSION	AR456323.1		GI:42691378
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unpublished.		
AUTHORS	1 (bases 1 to 2674)		
TITLE	Gai,W., Ye,J., Di Francesco,V. and Beasley,E.M.		
FEATURES	Isolated human kinase proteins, and uses thereof		
FEATURES	Patent: US 6686187-A 1 03-FEB-2004;		
source	Location/Qualifiers		
	1..2674		
ORIGIN	/organism="unknown"		
	/mol_type="genomic DNA"		
Query Match	100.0%;	Score 2674;	DB 6;
Beet Local Similarity	100.0%;	Pred. No. 0;	
Matches 2674;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps 0;
Qy	1	TCCGGGGTCCGCA	CGGGCTGAGTGGGTCCGAGGCGGTCCCGAGGACGAGCTGCCGTGC 60
Db	1	TCCGGGGTCCGCA	CGGGCTGAGTGGGTCCGAGGCGGTCCCGAGGACGAGCTGCCGTGC 60
Qy	61	GGAACAGCACTATGGGCTTCTTCTTGA	GTGTGAGCCGCCAGGGCCACGGGGTCTTGC 120
Db	61	GGAACAGCACTATGGGCTTCTTCTTGA	GTGTGAGCCGCCAGGGCCACGGGGTCTTGC 120
Qy	121	AGCAATTCAGAGAGGCCCGA	GCTTCTGTACTGAGAGGCATGAAAGTGAATGCCACGC 180
Db	121	AGCAATTCAGAGAGGCCCGA	GCTTCTGTACTGAGAGGCATGAAAGTGAATGCCACGC 180
Qy	181	GGGTCAAGATGACAGGAGATATG	CAGAGACTTCACCAATGTCCTTGACAGACATG 240
Db	181	GGGTCAAGATGACAGGAGATATG	CAGAGACTTCACCAATGTCCTTGACAGACATG 240
Qy	241	GGGGCAGAGCCGGGCA	TCAAGCCTTGACAGCCCATCAGTCACTGGGCTGAGATCA 300
Db	241	GGGGCAGAGCCGGGCA	TCAAGCCTTGACAGCCCATCAGTCACTGGGCTGAGATCA 300
Qy	301	CCAGCCAACTGAGGGCTTGA	GGCCCTTGCTGCGGACACGCAAGAGATCTGAATCAG 360
Db	301	CCAGCCAACTGAGGGCTTGA	GGCCCTTGCTGCGGACACGCAAGAGATCTGAATCAG 360
Qy	361	GGCCCTTGAAGCAAGCTGA	AGCTGTCAATCCGGGAACGGCAGACAGTTGGCAAGACTTACA 420
Db	361	GGCCCTTGAAGCAAGCTGA	AGCTGTCAATCCGGGAACGGCAGACAGTTGGCAAGACTTACA 420
Qy	421	CGAGCAATGGCAGCAAGCTTGA	CAGAGACTTCAACCAAGCCCAACGCAAGCAATTGAGA 480
Db	421	CGAGCAATGGCAGCAAGCTTGA	CAGAGACTTCAACCAAGCCCAACGCAAGCAATTGAGA 480

Db	422	GCAGAGCTGGCAGACACTGTGAGAGGAGCTCAACAAAGACCCACAGCCAGGACATTGTAGA	480
Qy	481	AGCTGAAGAGCCAGTACCAGACTCTGTGGACGGGACAGTGGCCAAAGCCAGGCAAGTACC	540
Db	481	AGCTGAAGAGCCAGTACCAGACTCTGTGGACGGGACAGTGGCCAAAGCCAGGCAAGTACC	540
Qy	541	AGGAGGCGACGAAGAACAAGGACCGTGAACAAGCCAAAGGACAATATATGTGGCAGCCGT	600
Db	541	AGGAGGCGACGAAGAACAAGGACCGTGAACAAGCCAAAGGACAATATATGTGGCAGCCGT	600
Qy	601	GGAAGCTCTTTGTCTCAACAACAACCGCTATGTGCTGGGCGTGGCGGCTGCGCAGTCAACC	660
Db	601	GGAAGCTCTTTGTCTCAACAACAACCGCTATGTGCTGGGCGTGGCGGCTGCGCAGTCAACC	660
Qy	661	ACCAGCACCAACCAACAGCTCTGCTGCGCCGCGCTGTGCGGTCACTGACAGACCTGCACG	720
Db	661	ACCAGCACCAACCAACAGCTCTGCTGCGCCGCGCTGTGCGGTCACTGACAGACCTGCACG	720
Qy	721	AGGAGATGGCTTGATCTTGAAGGAGATCTGTGAGGAATACCTGGAATTATGACAGCCGCG	780
Db	721	AGGAGATGGCTTGATCTTGAAGGAGATCTGTGAGGAATACCTGGAATTATGACAGCCGCG	780
Qy	781	TGCAGAGATGAGAGTGTGTGGCCATTACACGGGAGATGGCTGACGCTGTGCCGCACTCAGC	840
Db	781	TGCAGAGATGAGAGTGTGTGGCCATTACACGGGAGATGGCTGACGCTGTGCCGCACTCAGC	840
Qy	841	CTGAGGCTGAGTACCAAGGCTTCTGTGCAAGATATGGGTCCGCACTGACGTTCCACCTT	900
Db	841	CTGAGGCTGAGTACCAAGGCTTCTGTGCAAGATATGGGTCCGCACTGACGTTCCACCTT	900
Qy	901	GTGTCAAGTTTCGATGAGTCACTGCTTGAAGAGGATGAACCGCTGAGCTTGTGGGAGCTCC	960
Db	901	GTGTCAAGTTTCGATGAGTCACTGCTTGAAGAGGATGAACCGCTGAGCTTGTGGGAGCTCC	960
Qy	961	AGCTGAACGAGCTGACTGTGGAGACGTGCAGCAACGCTGACCTCAATGACATGATGAC	1020
Db	961	AGCTGAACGAGCTGACTGTGGAGACGTGCAGCAACGCTGACCTCAATGACATGATGAC	1020
Qy	1021	TGGCTGTGGCCACCGAAGTGGTGTTCAGCGCGGACAGAGATGGTTATCGACGCTGCACACG	1080
Db	1021	TGGCTGTGGCCACCGAAGTGGTGTTCAGCGCGGACAGAGATGGTTATCGACGCTGCACACG	1080
Qy	1081	AGCTCCGGAATGAAGAGAGAAACAACCCACCCCGGGAGCCGGGTCACTGTGGGCAAGA	1140
Db	1081	AGCTCCGGAATGAAGAGAGAAACAACCCACCCCGGGAGCCGGGTCACTGTGGGCAAGA	1140
Qy	1141	GGCAAGTGTCCCAAGAGCACTGCAAGGGGCTGCAGGATAGGCTGTGACGCCAGGCCAAGC	1200
Db	1141	GGCAAGTGTCCCAAGAGCACTGCAAGGGGCTGCAGGATAGGCTGTGACGCCAGGCCAAGC	1200
Qy	1201	TGCAGAGCCCAAGAGATTTGTCTGACAGCAAAAGCTGTGAGCACTTGGGCCCGCGACGCC	1260
Db	1201	TGCAGAGCCCAAGAGATTTGTCTGACAGCAAAAGCTGTGAGCACTTGGGCCCGCGACGCC	1260
Qy	1261	CGCTGTGTCTCTCTGTGAGGATGACCGCACTTCACGTGTGCTCTCGAGCAGAGCGAG	1320
Db	1261	CGCTGTGTCTCTCTGTGAGGATGACCGCACTTCACGTGTGCTCTCGAGCAGAGCGAG	1320
Qy	1321	AGGGGGGGAAGACACCAACGCTGTGAATCTTTAAGACCAATTCACAGGAATCTTTCCGCC	1380
Db	1321	AGGGGGGGAAGACACCAACGCTGTGAATCTTTAAGACCAATTCACAGGAATCTTTCCGCC	1380
Qy	1381	CCAAGTTCTCGAATCTGTACCGCATGTGAAGGGAGAGGCTTTCTTAGCATTCCTTTGTCTCA	1440
Db	1381	CCAAGTTCTCGAATCTGTACCGCATGTGAAGGGAGAGGCTTTCTTAGCATTCCTTTGTCTCA	1440
Qy	1441	TCGACCAACTATCTGAGCAACCAAGACGCCCTTCAACAAGAGAGTGATGTGTCTTGCACACA	1500
Db	1441	TCGACCAACTATCTGAGCAACCAAGACGCCCTTCAACAAGAGAGTGATGTGTCTTGCACACA	1500
Qy	1501	GGGCTGTGTCCCAAGAGCAATGTGGTGTCTGAACATATGAGGACCTGTGTGTGGGTATGACGA	1560
Db	1501	GGGCTGTGTCCCAAGAGCAATGTGGTGTCTGAACATATGAGGACCTGTGTGTGGGTATGACGA	1560

/feature type="Brain, adult medulla"
/clone_id="NH1.MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
1. 2788
/gene="FBS"
/note="synonym: FBS"
/db_xref="LOCUSID:2242"
/db_xref="MIM:190030"
73. 2541
/gene="FBS"
/codon_start=1
/product="V-FBS, feline sarcoma viral/V-fps fujinami avian
sarcoma viral oncogene homolog"
/protein_id="AAH35357.1"
/db_xref="GI:23271525"
/db_xref="LOCUSID:2242"
/db_xref="MIM:190030"
/translation="MGFSSSELSPPGSHGVLLQOMQAEALRLLEGMRKMAORVKSREY
AGLHHMSIQDSGGSRATSPDSPTQSHAEITSTQEGSLRLRQAEHLNKGPSKL
SLIRRQQLRTYSQWQQLQDELTKTSHSDI EKUKSYRALASDAKKRYDEAS
KOKDRKADKTVRSIMKLFAHNRVVLGVRRAQAHQHHQLLPGLRLSDLHEE
MACILKEILOEYLEISLVQDEVVAIHRMAAARIOBEAYGFLROYGSAADVP
CYTDFESLIEBGPFLPELOINELTVESVOTLTSTDELAVATMVFROEMTOL
QOELNREENTHPRERVOILGKROVALQGLQVALGSAQLAQOELLOTKEHIG
PBBPVLILQDDRHSTSSBQREBGRPTTLEILKSHSGIFRPFPSLPPLQILPE
VPRPHEQLWYHGATPRAEVALVHSGDFLVRESQKQEVLSVLMDELPHFIIOQ
LNLVRLIEGEPSPILIDHLSTQPLTKSGVLAAPDKVNLNHEBVLVLEQ
IRGNGEVPSGLRADNTLVAVKSCEFTPLPKAKFLQEARILKOYBHPVIRLIG
VCTOKPIYIVMELVGGDFLFLRTGEGARLRYKTLLOWGDAALMEYKSCCHIR
DLAARCLVTEKNVLKISDRCMSREBADVAAAGGLRVPVKMTAPALNTGRASSE
SPVMSFGILMETFSLGASPYRPLSNQOTREVEKGRILPCEBELCDVAFRLMEQMA
YEPGQPSFTIYQELQSIKQHR"

ORIGIN

Query Match 88.1%; Score 2356; DB 9; Length 2788;
Best Local Similarity 92.5%; Pred. No. 0;

Matches 2576; Conservative 0; Mismatches 0; Indels 210; Gaps 1;

QY 2 CCGGGGTCGCGACCGGGCTGAGTCGCTCGAGCGCGTCCGAGAGCAGCTGCGCGG 61
DB 3 CCGGGGTCGCGACCGGGCTGAGTCGCTCGAGCGCGTCCGAGAGCAGCTGCGCGG 62
QY 62 GAAACAGACTTATGGCTTCTCTTCTGAGCTGTGCAAGCCCGCAGGGGCAAGGGTCTTGA 121
DB 63 GAAACAGACTTATGGCTTCTCTTCTGAGCTGTGCAAGCCCGCAGGGGCAAGGGTCTTGA 122
QY 122 GCAAAATGCAAGAGCGCGAGCTTCTCTAATGAGGGGATGAGAAAGTGAATGAGCGGAG 181
DB 123 GCAAAATGCAAGAGCGCGAGCTTCTCTAATGAGGGGATGAGAAAGTGAATGAGCGGAG 182
QY 182 GGTCAAGAGTGAAGAGAGTATGACAGAGTGTCTTCAACATGTCTCTGAGAGACAGTGA 241
DB 183 GGTCAAGAGTGAAGAGAGTATGACAGAGTGTCTTCAACATGTCTCTGAGAGACAGTGA 242
QY 242 GGGCCAGAGCGGGGCAATGAGCTTCAAGAGCCCATCATGTCAGTCTTGGCTGATGATC 301
DB 243 GGGCCAGAGCGGGGCAATGAGCTTCAAGAGCCCATCATGTCAGTCTTGGCTGATGATC 302
QY 302 GAGCAAACTGAAGGGGCTGAGCGGCTTGTGTCGGGGGAGAGCGAGAGATCTGAATCTGAG 361
DB 303 GAGCAAACTGAAGGGGCTGAGCGGCTTGTGTCGGGGGAGAGCGAGAGATCTGAATCTGAG 362
QY 362 GCGCTTGAAGAGTGAAGCTGTCTCATCCGGGAAACGGCAGACAGCTTTCGCAAGACTTACAG 421
DB 363 GCGCTTGAAGAGTGAAGCTGTCTCATCCGGGAAACGGCAGACAGCTTTCGCAAGACTTACAG 422
QY 422 CGAGCAGTGGCAGACACTGACAGAGAGTCTACCAAGACCCACAGAGCAGAGCACTTACAGAA 481
DB 423 CGAGCAGTGGCAGACACTGACAGAGAGTCTACCAAGACCCACAGAGCAGAGCACTTACAGAA 482
QY 482 GCTGAAGAGCAGTACAGAGCTTGGCAGAGGACAGTGGCCAAAGCCAAAGGACAGTACCA 541
DB 483 GCTGAAGAGCAGTACAGAGCTTGGCAGAGGACAGTGGCCAAAGCCAAAGGACAGTACCA 542

DB 483 GCTGAAGAGCAGTACAGAGCTTGGCAGAGGACAGTGGCCAAAGCCAAAGGACAGTACCA 542
QY 542 GGAAGCCAGACAAAGACAGAGACCTGTGAACAAGCCAAAGACAAAGTATGTGGCAGCTTGA 601
DB 543 GGAAGCCAGACAAAGACAGAGACCTGTGAACAAGCCAAAGACAAAGTATGTGGCAGCTTGA 602
QY 602 GAAAGCTCTTGTGTCACCAACACCGGATGTGTGTCGGGGCGGGGGCTGGCAGCTTACCA 661
DB 603 GAAAGCTCTTGTGTCACCAACACCGGATGTGTGTCGGGGCGGGGGCTGGCAGCTTACCA 662
QY 662 CCAGCACAACCAACAGCTCTGCTGCGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
DB 663 CCAGCACAACCAACAGCTCTGCTGCGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 722
QY 722 GGAAGTGGCTTGCATCTGAAGAGATCTGCAAGAAATCTTGAAGATTAGCAGCTTGGT 781
DB 723 GGAAGTGGCTTGCATCTGAAGAGATCTGCAAGAAATCTTGAAGATTAGCAGCTTGGT 782
QY 782 GGAAGTGGCTTGCATCTGAAGAGATCTGCAAGAAATCTTGAAGATTAGCAGCTTGGT 841
DB 783 GGAAGTGGCTTGCATCTGAAGAGATCTGCAAGAAATCTTGAAGATTAGCAGCTTGGT 842
QY 842 TGAGGCTGAGTACCAAGCTTCTGCGACAGTATGGGTCCGACCTGACCTGACCTG 901
DB 843 TGAGGCTGAGTACCAAGCTTCTGCGACAGTATGGGTCCGACCTGACCTGACCTG 902
QY 902 TGTCACTTTCATGATGATCACTGCTTGAAGAGGATGAACCTGAGCTTGGGAGCTTCA 961
DB 903 TGTCACTTTCATGATGATCACTGCTTGAAGAGGATGAACCTGAGCTTGGGAGCTTCA 962
QY 962 GCTGAACGAGCTGACCTGTGAGAGAGGCTGAGACACGCTGACCTGACCTGACCTGAGCT 1021
DB 963 GCTGAACGAGCTGACCTGTGAGAGAGGCTGAGACACGCTGACCTGACCTGACCTGAGCT 1022
QY 1022 GCTGTGGCCACCGAGATGCTTCAAGCGGAGAGAGATGTTACGACCTGACCAAGAG 1081
DB 1023 GCTGTGGCCACCGAGATGCTTCAAGCGGAGAGAGATGTTACGACCTGACCAAGAG 1082
QY 1082 GCTTCCGAAATGAAGAGAGACACCAACCCCGGGAGCGGGTGCAGCTGTGGGCAAGAG 1141
DB 1083 GCTTCCGAAATGAAGAGAGACACCAACCCCGGGAGCGGGTGCAGCTGTGGGCAAGAG 1142
QY 1142 GCAAGTGTGCAAGAGACCTGCAAGGGGCTGAGAGGAGGCTGTGAGCAGAGCCAAAGCT 1201
DB 1202 GCAAGTGTGCAAGAGACCTGCAAGGGGCTGAGAGGAGGAGGCTGTGAGCAGAGCCAAAGCT 1202
QY 1202 GCAAGTGTGCAAGAGACCTGCAAGGGGCTGAGAGGAGGAGGCTGTGAGCAGAGCCAAAGCT 1202
DB 1203 GCAAGTGTGCAAGAGACCTGCAAGGGGCTGAGAGGAGGAGGCTGTGAGCAGAGCCAAAGCT 1202
QY 1262 GCTGTGTGCTTCTTGAAGATGACCGGCACTTCAAGTGTCTTCTGAGAGAGAGCTGAG 1321
DB 1263 GCTGTGTGCTTCTTGAAGATGACCGGCACTTCAAGTGTCTTCTGAGAGAGAGCTGAG 1322
QY 1322 GGGGGGAAGAACCAACCGAGCTGAGATCTTGAAGAGCCATCTGAGAAATCTTGGGCC 1381
DB 1323 GGGGGGAAGAACCAACCGAGCTGAGATCTTGAAGAGCCATCTGAGAAATCTTGGGCC 1382
QY 1382 CAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1391
DB 1383 CAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1391
QY 1392 GCAAGTGTGTAACAAGGGGCACTCCGAGGGCAGAGGTGCTGAGCTGTGGTGCATCTC 1502
DB 1443 GCAAGTGTGTAACAAGGGGCACTCCGAGGGCAGAGGTGCTGAGCTGTGGTGCATCTC 1502
QY 1392 GGGGAGCTTCTGAGTGGGAGAGCCAGGGCAAGAGAGTACGTGCTGTGGTGTGG 1562
DB 1503 TGGGAGCTTCTGAGTGGGAGAGCCAGGGCAAGAGAGTACGTGCTGTGGTGTGG 1562
QY 1392 GGGGAGCTTCTGAGTGGGAGAGCCAGGGCAAGAGAGTACGTGCTGTGGTGTGG 1562
DB 1392 GGGGAGCTTCTGAGTGGGAGAGCCAGGGCAAGAGAGTACGTGCTGTGGTGTGG 1562
QY 1392 GGGGAGCTTCTGAGTGGGAGAGCCAGGGCAAGAGAGTACGTGCTGTGGTGTGG 1562
DB 1563 GGAATGTCTGCGCCGCACTTGCATCATCAAGTCTTGAATTAACCTGTATCAAGCTGGAAG 1622

Qy	122	GCAAATGCAAGAGGCGGAGCTTCGTCATCTGAGAGGCAATGAGAAATGATGATGCGCCACG	181
Dp	126	GCAATATGAGAGAGGCGGAGCTTCGTCATCTGAGAGGCAATGAGAAATGATGATGCGCCACG	185
Qy	182	GGTCAAGAGTACACAGGAGATATGACAGACTGCTTACCAATGTCTCTGACAGGACATGTG	241
Dp	186	GGTCAAGAGTACACAGGAGATATGACAGACTGCTTACCAATGTCTCTGACAGGACATGTG	245
Qy	242	GGGCCAGAGCCGGGCAATCAGCCTGACAGCCCATCACTCATGTCCTGTGGCTGATATAC	301
Dp	246	GGGCCAGAGCCGGGCAATCAGCCTGACAGCCCATCACTCATGTCCTGTGGCTGATATAC	305
Qy	302	CAGCCAACTBAGGGGCTGAGCCGCTTGTCTGCGAGAGACCCAGAGAGATCTGAATCTCAG	361
Dp	306	CAGCCAACTBAGGGGCTGAGCCGCTTGTCTGCGAGAGACCCAGAGAGATCTGAATCTCAG	365
Qy	362	GGCCCTGAGCAAGCTGAGCCTGCTCATCTCGGGAAACGGAGAGACTTGCACAAGACTTACAG	421
Dp	366	GGCCCTGAGCAAGCTGAGCCTGCTCATCTCGGGAAACGGAGAGACTTGCACAAGACTTACAG	425
Qy	422	CGAGCAGTGGCAGCAGCTGCGACAGAGAGCTCACCAAGACCAACAGCCAGACATTTGAGA	481
Dp	426	CGAGCAGTGGCAGCAGCTGCGACAGAGAGCTCACCAAGACCAACAGCCAGACATTTGAGA	485
Qy	482	GCTGAAGAGCCAGTATCCGAGCTCTGCGACGGGACAGTGGCCCAAGCCAGCCAGTATCCA	541
Dp	486	GCTGAAGAGCCAGTATCCGAGCTCTGCGACGGGACAGTGGCCCAAGCCAGCCAGTATCCA	545
Qy	542	GAGGCGCAGCAAGAGCAAGGACCCGTGACAAAGGCCAAAGGACAAATATGATGGCAGGCTGTG	601
Dp	546	GAGGCGCAGCAAGAGCAAGGACCCGTGACAAAGGCCAAAGGACAAATATGATGGCAGGCTGTG	605
Qy	602	GAACTCTTTGCTCAACCAACCCGCTATGTGTGGGCGTGGGCGTGGCGCACTTACACA	661
Dp	606	GAACTCTTTGCTCAACCAACCCGCTATGTGTGGGCGTGGGCGTGGCGCACTTACACA	665
Qy	662	CCAGCACCAACCAACGAGCTCTGTCTGCCGGGCTGCTGCGGTCTACAGACACTGACAGA	721
Dp	666	CCAGCACCAACCAACGAGCTCTGTCTGCCGGGCTGCTGCGGTCTACAGACACTGACAGA	725
Qy	722	GGAGTGGCTTGCATCTCGAAGAGAGATCTCTCAGAGGAATACTGGAGATTACAGCTGGT	781
Dp	726	GGAGTGGCTTGCATCTCGAAGAGAGATCTCTCAGAGGAATACTGGAGATTACAGCTGGT	785
Qy	782	GCAGGATGAGGTGTGTGGCATTTCAACCGGAGATGTGCTCAGCTGTGCCGATCTCAGCC	841
Dp	786	GCAGGATGAGGTGTGTGGCATTTCAACCGGAGATGTGCTCAGCTGTGCCGATCTCAGCC	845
Qy	842	TGAGGCTGATGATCCAAAGGCTTCTGTGCAAGATATGGATCCGACCTTGACGTCCACCTGT	901
Dp	846	TGAGGCTGATGATCCAAAGGCTTCTGTGCAAGATATGGATCCGACCTTGACGTCCACCTGT	905
Qy	902	TGTACACGTTGCATGATGATCATGCTTTGABAGAGGTGAACCGTGTGAGCCTTGGGGAGCTTCA	961
Dp	906	TGTACACGTTGCATGATGATCATGCTTTGABAGAGGTGAACCGTGTGAGCCTTGGGGAGCTTCA	965
Qy	962	GCTGAACAGAGCTGACTGTGAGAGAGGCTGACACAACGCTGACCTTACATGACAGATGAGCT	1021
Dp	966	GCTGAACAGAGCTGACTGTGAGAGAGGCTGACACAACGCTGACCTTACATGACAGATGAGCT	1025
Qy	1022	GGCTGTGGCCACCGAGATGTGTTCACGGCGGCAAGAGATGTGTTACGACCTGCAACAGGA	1081
Dp	1026	GGCTGTGGCCACCGAGATGTGTTCACGGCGGCAAGAGATGTGTTACGACCTGCAACAGGA	1085
Qy	1082	GCTCCGGAATGAATAGAGAGAACCCCAACCCCGGGAGCGGGGTGCAAGCTGTGGGCAAGAG	1141
Dp	1086	GCTCCGGAATGAATAGAGAGAACCCCAACCCCGGGAGCGGGGTGCAAGCTGTGGGCAAGAG	1145
Qy	1142	GCAAGTGTGCAAGAGACATGCAAGGGGCTGACAGGTACGCTGTGACGACCAAGCCAACT	1201
Dp	1146	GCAAGTGTGCAAGAGACATGCAAGGGGCTGACAGGTACGCTGTGACGACCAAGCCAACT	1205
Qy	1202	GCAAGCCCAAGAGAGTTGCTTGACAGACCAAGCTGGAGCACTTGGGCGCCGCGAGGCCCC	1261

Db	1206	GCAGGCCAGAGAGATTGCTGCAAGCCAACTGAGCACTTGGGCCCCGGCGAGCCCC	1265
Oy	1262	GCCTGTGCTGCTCTGCAAGATGACCGGCATCTCGACGTGCTCTTCGAGCAGAGACGAGA	1321
Db	1265	GCTGTGTGCTGCTCTCGAGAGATGACCGGCATCTCCACGTCGTCTTCGAGCAGAGACGAGA	1322
Oy	1332	GGGGGAAAGACATCCACGCTGGAGATCTTAAAGCCACATCTCAGGAATCTTCCGCC	1381
Db	1336	GGGGGAAAGGACATCCACGCTGGAGATCTTAAAGCCACATCTCAGGAATCTTCCGCC	1385
Oy	1382	CAAGTCTCG-----	1391
Db	1386	CAAGTCTCGCTCCCTCCACCGCTGACGCTCATTCGGAGGTGCAAGAGCCCTGCAATGA	1441
Oy	1392	-----	1391
Db	1446	GCACGTGATGACACACGGGGGCATTCGAGGGGCAAGAGTGGCTAGAGCTGCTGATGCACTC	1505
Oy	1392	-----	1391
Db	1506	TGGGGACTTCCTGGTGGCGAGAGCCAGGGCAACAGAGATGACTGTGCTGGTCTGTG	1565
Oy	1392	-----	1391
Db	1566	GGATGCTTCGCCCCGGGCATTCATCATTCAGTCTTGATTAACCTGTACCGACTGGAAAG	1625
Oy	1412	GGAAAGCTTTCTTAGCATCTTTGCTCATGACCACTACTGAGCAACCCAGCAGCCCT	1471
Db	1636	GGAAAGCTTTCTTAGCATCTTTGCTCATGACCACTACTGAGCAACCCAGCAGCCCT	1685
Oy	1472	CACCAAGAAGTGTGTGTCCTGCAACAGGGCTGTGGCCCAAGACAAAGTGGTCTGAA	1531
Db	1686	CACCAAGAAGTGTGTGTCCTGCAACAGGGCTGTGGCCCAAGACAAAGTGGTCTGAA	1745
Oy	1532	CCATGAGGACCTGCTGTGGTGTGAGCAGATTGACACGGGGAACTTTGGCCAAAGTCTCAG	1591
Db	1746	CCATGAGGACCTGCTGTGGTGTGAGCAGATTGACACGGGGAACTTTGGCCAAAGTCTCAG	1805
Oy	1592	CGGACGCTGGGAGCCGCAACAACCTGGTGGCGGTGAAGTCTTGTGAGAGACGCTCC	1651
Db	1806	CGGACGCTGGGAGCCGCAACAACCTGGTGGCGGTGAAGTCTTGTGAGAGACGCTCC	1865
Oy	1652	ACCTGACCTCAAGGCCAAAGTTTCTACAGAAACGAGGATCCTGAAGCAGTACAGCAACC	1711
Db	1866	ACCTGACCTCAAGGCCAAAGTTTCTACAGAAACGAGGATCCTGAAGCAGTACAGCAACC	1925
Oy	1712	CAACATGTGTGCTCATTTGTTGTCGACCCAGAAAGAGGCCATCTACATGCTCATGGA	1771
Db	1926	CAACATGTGTGCTCATTTGTTGTTGTCGACCCAGAAAGAGGCCATCTACATGCTCATGGA	1985
Oy	1772	GCTTGTGAGGGGGGCGACTTCGAGCCTTCCTCGGACCGAGGGGGGCGGCTGCGGGT	1831
Db	1986	GCTTGTGAGGGGGGCGACTTCGAGCCTTCCTCGGACCGAGGGGGGCGGCTGCGGGT	2045
Oy	1832	GAAAGACTCTGTGCAAGATGTGTGGGGATGACAGTCTGGCAATGAGATACCTGAGAGCAA	1891
Db	2046	GAAAGACTCTGTGCAAGATGTGTGGGGATGACAGTCTGGCAATGAGATACCTGAGAGCAA	2105
Oy	1892	GTCGTGATTCACCGGGACTGTGCTGCTCGGAATGCTGCTGTGACAGAGAAATGTCCT	1951
Db	2106	GTCGTGATTCACCGGGACTGTGCTGCTCGGAATGCTGCTGTGACAGAGAAATGTCCT	2165
Oy	1952	GAAATCATGATGACTTTGGGATGTCCCGAGAGAAAGCGATGGGGTCTATGCAAGCTCAGG	2011
Db	2166	GAAATCATGATGACTTTGGGATGTCCCGAGAGAAAGCGATGGGGTCTATGCAAGCTCAGG	2225
Oy	2012	GGGGCTCAGACAAATCCCGTGAAGTGAACGCACTGAGGCTCTTAATCTACGCGCGTA	2071
Db	2226	GGGGCTCAGACAAATCCCGTGAAGTGAACGCACTGAGGCTCTTAATCTACGCGCGTA	2285
Oy	2072	CTCCTCCGAAAGCAAGTGTGAGACTTTGCACTTGTCTGAGAGACTTCAAGCTTGGG	2131

Db 2286 CTCCTCCGAAGCGAGCTGTGAGCTTTGGCATCTTGTCTGTGGAGACCTTCAGCCCTGG 2345
QY 2132 GGGCTCCCTCTATCCCAACCTGACATGACGAGACCGGGAGTTGTGAGAGAGGGGGG 2191
Db 2346 GGGCTCCCTCTATCCCAACCTGACATGACGAGACCGGGAGTTGTGAGAGAGGGGGG 2405
QY 2192 CCGTCTGCCCTGCGCAGAGCTGTGTCTGTGAGCCGTTCAGAGCTCATGAGAGAGTGTG 2251
Db 2406 CCGTCTGCCCTGCGCAGAGCTGTGTCTGTGAGCCGTTCAGAGCTCATGAGAGAGTGTG 2465
QY 2252 GGGCTGTAGAGCTGTGGAGAGCGGCCAGCTTCAGACACCATTTACAGAGAGCTGACAGCAT 2311
Db 2466 GGGCTGTAGAGCTGTGGAGAGCGGCCAGCTTCAGACACCATTTACAGAGAGCTGACAGCAT 2525
QY 2312 CGAAGAGCGGAGCTGGTGGAGGCTGGGAGCCCTTCAGAGCTGTGGCCCTGTCAGAGCCCT 2371
Db 2526 CGAAGAGCGGAGCTGGTGGAGGCTGGGAGCCCTTCAGAGCTGTGGCCCTGTCAGAGCCCT 2585
QY 2372 AGGTGAGAGCTCTCAGAGGCTCCAGCTCATATGCTGACAGCTCTTCAAGTCTTGAGACTC 2431
Db 2586 AGGTGAGAGCTCTCAGAGGCTCCAGGCTCATATGCTGACAGCTCTTCAAGTCTTGAGACTC 2645
QY 2432 CTGCCACAGCATCCACACTGCGCGAGATGACAGCGCGGTCTCTCTGTGTCCCTGC 2491
Db 2646 CTGCCACAGCATCCACACTGCGCGAGATGACAGCGCGGTCTCTCTGTGTCCCTGC 2705
QY 2492 TGGTGCAGGAGCTCTCTCTTCCGGGAGAAACAATTAACACTTGTGCCACTGAAAAA 2551
Db 2706 TGGTGCAGGAGCTCTCTCTTCCGGGAGAAACAATTAACACTTGTGCCACTGAAAAA 2765
QY 2552 AAAAAA 2557
Db 2766 AAAAAA 2771

RESULT 5
AX281847 2889 bp DNA linear PAT 02-NOV-2001
LOCUS AX281847
DEFINITION Sequence 256 from Patent WO0177389.
ACCESSION AX281847
VERSION AX281847.1 GI:16609098
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 Shiffman, D., Somogyi, R., Lawn, R., Seilhamer, J.J., Porter, G.J.,
Mikita, T. and Tai, J.
TITLE Patent expressed in foam cell differentiation
JOURNAL Patent: WO 0177389-A 256 18-OCT-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1. .2889
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 997347.6"

ORIGIN
Query Match 87.0%; Score 2326.4; DB 6; Length 2889;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 2558; Conservative 0; Mismatches 1; Indels 211; Gaps 2;

Db 2238 GCAATGACAGAGCGGAGCTTCGTTACTGTGAGGAGCAAGAAAGTGTAGTGGCCAGCG 297
QY 182 GGTCAAGAGTGA CAGGAGTATGAGAGACTGTTCACCAAGTCCCTGACAGAGCTGG 241
Db 298 GGTCAAGAGTGA CAGGAGTATGAGAGACTGTTCACCAAGTCCCTGACAGAGCTGG 357
QY 242 GGGCAGAGCGGGGACATGAGCCCTGACAGCCCATCATGATCAGTCTGTGGCTGAGATCAG 301
Db 358 GGGCAGAGCGGGGACATGAGCCCTGACAGCCCATCATGATCAGTCTGTGGCTGAGATCAG 417
QY 302 CAGCCAACTGA -GGGCTGAGCCCTGTCTGCGGAGACGAGAGATCTGAACTCAG 360
Db 418 CAGCCAACTGAGGGGCTGAGCCCTGTCTGCGGAGACGAGAGATCTGAACTCAG 477
QY 361 GGGCCCTGAGCAAGCTGAGCCCTGATCCCGGGAACCGGAGAGCTTGGCAAGACTTACA 420
Db 478 GGGCCCTGAGCAAGCTGAGCCCTGTCTATCCGGGAACCGGAGAGCTTGGCAAGACTTACA 537
QY 421 GCGAGAGTGGAGAGAGCTGACAGAGAGCTCACCAAGACCCAGACAGCATTTAGA 480
Db 538 GCGAGAGTGGAGAGAGCTGACAGAGAGCTCACCAAGACCCAGACAGCATTTAGA 597
QY 481 AGCTGAAGAGCAAGTACCGAGCTTGGCAAGGAGCTCCCAAGCCAAAGCCAAAGTACC 540
Db 598 AGCTGAAGAGCAAGTACCGAGCTTGGCAAGGAGCTCCCAAGCCAAAGCCAAAGTACC 657
QY 541 AGGAGGCGGAGCAAGACAAAGAGCCGTGACAAAGGCAAGGACAAAGTATGGGAGCTGT 600
Db 658 AGGAGGCGGAGCAAGACAAAGAGCCGTGACAAAGGCAAGGACAAAGTATGGGAGCTGT 717
QY 601 GGAAGCTTTTGTCTGACCAACAACCGCTATGTCTGGGCGTGGGAGCTGACAGCTTACAC 660
Db 718 GGAAGCTTTTGTCTGACCAACAACCGCTATGTCTGGGCGTGGGAGCTTACAC 777
QY 661 ACCAGACCAACCAACAGCTCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 778 ACCAGACCAACCAACAGCTCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837
QY 721 AGGAGATGGCTTGCATCTCGAAGAGAGATCTCGAGGAATACCTGAGATTAGCAGCTGG 780
Db 838 AGGAGATGGCTTGCATCTCGAAGAGAGATCTCGAGGAATACCTGAGATTAGCAGCTGG 897
QY 781 TCAGAGATGAGTGTGGGCTCATTCACCGGAGATGAGCTGAGCTGCGCGCATCAGC 840
Db 898 TCAGAGATGAGTGTGGGCTCATTCACCGGAGATGAGCTGAGCTGCGCGCATCAGC 957
QY 841 CTGAGGCTGAGTACCAAGGCTTCTGCGACAGTATGGTCCGACCTGACGCTCCACCT 900
Db 958 CTGAGGCTGAGTACCAAGGCTTCTGCGACAGTATGGTCCGACCTGACGCTCCACCT 1017
QY 901 GGTGCAAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
Db 1018 GGTGCAAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1077
QY 961 ACTGAACAGAGTGA CTGTGAGAGAGGTGACACACAGCTGACCTTCAAGTATGACAGTAC 1020
Db 1078 ACTGAACAGAGTGA CTGTGAGAGAGGTGACACACAGCTGACCTTCAAGTATGACAGTAC 1137
QY 1021 TGGCTGTGCGCACCGAGATGTGTTTCAAGCGGAGAGATGTTTACGACGCTGCAACAG 1080
Db 1138 TGGCTGTGCGCACCGAGATGTGTTTCAAGCGGAGAGATGTTTACGACGCTGCAACAG 1197
QY 1081 AGCTCCGGAATGAAGAGAGAAACCCACCCCGGAGAGGAGGAGGAGGAGGAGGAGGAGG 1140
Db 1198 AGCTCCGGAATGAAGAGAGAAACCCACCCCGGAGAGGAGGAGGAGGAGGAGGAGGAGG 1257
QY 1141 GCGAAGTGTGCAAGAGACATGACAGGAGCTGACAGGATGAGCTGTGACAGGAGGAGGAGG 1200
Db 1258 GCGAAGTGTGCAAGAGACATGACAGGAGCTGACAGGATGAGCTGTGACAGGAGGAGGAGG 1317
QY 1201 TGCAGGCGGAGAGAGGTTGCTGACAGACCAAGCTGAGCACTTGGGCGGAGGAGGAGGAGG 1260

Db 1318 TGCAGGCCAGAGAGATTGCTGCAGACCAAGCTGGAGCACTGGGCCCCGGGAGAGCCCC 1377
 Qy 1261 CGCTTGTCTGCTCTCTGCAGAGATGACCGCACTCCAGTGTCTCTGCAGAGAGAGAGAG 1320
 Db 1378 CGCTGTGTCTGCTCTCTGCAGAGATGACCGCACTCCAGTGTCTCTGCAGAGAGAGAGAG 1437
 Qy 1321 AGGGGGGAGAGAGACCCAGCGCTGGAGATCTTTAAGAGCAACCTCTAGAGAACTTTCCGCGC 1380
 Db 1438 AGGGGGGAGAGAGACCCAGCGCTGGAGATCTTTAAGAGCAACCTCTAGAGAACTTTCCGCGC 1497
 Qy 1381 CCAAGTTCTG----- 1391
 Db 1498 CCAAGTTCTGCGTCCCTCCAGCGCTGACGTATTCGAGAGAGAGAGAGAGAGAGAGAGAG 1557
 Qy 1392 ----- 1391
 Db 1558 AGCAGCTGTGTACCAAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1617
 Qy 1392 ----- 1391
 Db 1618 CTGGGAGCTTCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1677
 Qy 1392 -----AACCTGTACCGAGTGAAG 1410
 Db 1678 GGGATGTCTGCGCGGAG 1737
 Qy 1411 GGGAGAGCTTCTCTGAGATCTCTTGTCTGATGACCACTGAGAGAGAGAGAGAGAGAGAG 1470
 Db 1738 GGGAGAGCTTCTCTGAGATCTCTTGTCTGATGACCACTGAGAGAGAGAGAGAGAGAGAG 1797
 Qy 1471 TCACCAAG 1530
 Db 1798 TCACCAAG 1857
 Qy 1531 ACCATGAG 1590
 Db 1858 ACCATGAG 1917
 Qy 1591 GGGAGAGCTTCTGAG 1650
 Db 1918 GGGAGAGCTTCTGAG 1977
 Qy 1651 GAGCTGAG 1710
 Db 1978 GAGCTGAG 2037
 Qy 1711 CCAAGATCGTGTCTCATTTGATGTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1770
 Db 2038 CCAAGATCGTGTCTCATTTGATGTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2097
 Qy 1771 AGCTTGTGAG 1830
 Db 2098 AGCTTGTGAG 2157
 Qy 1831 TGAAGATCTGCTGAG 1890
 Db 2158 TGAAGATCTGCTGAG 2217
 Qy 1891 AGTGTGATCTGACAG 1950
 Db 2218 AGTGTGATCTGACAG 2277
 Qy 1951 TGAAGATCTGATCTTTGGAGATGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2010
 Db 2278 TGAAGATCTGATCTTTGGAGATGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2337
 Qy 2011 GGGGGCTGAG 2070
 Db 2338 GGGGGCTGAG 2397
 Qy 2071 ACTCTCTCGAAG 2130
 Db 2398 ACTCTCTCGAAG 2457

Qy 2131 GGGCTCTCCCTCTATCCCACTTCAGCATCAGCAGACAGAGAGAGAGAGAGAGAGAGAGAG 2190
 Db 2458 GGGCTCTCCCTCTATCCCACTTCAGCATCAGCAGACAGAGAGAGAGAGAGAGAGAGAGAG 2517
 Qy 2191 GCGGTCTGCTGCGGAG 2250
 Db 2518 GCGGTCTGCTGCGGAG 2577
 Qy 2251 GGGCTCTATGAGCTGTGGAG 2310
 Db 2578 GGGCTCTATGAGCTGTGGAG 2637
 Qy 2311 TCCGAAGCGGATCGGTGAG 2370
 Db 2638 TCCGAAGCGGATCGGTGAG 2697
 Qy 2371 TAGGTGAGCTCTGAG 2430
 Db 2698 TAGGTGAGCTCTGAG 2757
 Qy 2431 CCGGACAGAGATCCACAGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2490
 Db 2758 CCGGACAGAGATCCACAGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2817
 Qy 2491 CTGCTGCGAGAGCTTCTCTTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2550
 Db 2818 CTGCTGCGAGAGCTTCTCTTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2877
 Qy 2551 AAAAAAAAAA 2560
 Db 2878 AAAAAAAAAA 2887

RESULT 6
 LOCUS CQ725220 2697 bp DNA linear PART 03-FEB-2004
 DEFINITION Sequence 11154 from Patent WO02068579.
 ACCESSION CQ725220
 VERSION CQ725220.1 GI:42286077
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 TITLE Kites, such as nucleic acid arrays, comprising a majority of
 humenexons or transcripts, for detecting expression and other uses
 thereof
 JOURNAL Patent: WO 02068579-A 11154 06-SEP-2002;
 PE Corporation (NY) (US)
 FEATURES
 source 1..2697
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 84.8%; Score 2267; DB 6; Length 2697;
 Best Local Similarity 92.2%; Pred. No. 0; Mismatches 0; Indels 210; Gaps 1;
 Matches 2487; Conservative 0;

Db 62 GAACAGCACTATGAGCTTCTCTTCTGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 121
 1 GAACAGCACTATGAGCTTCTCTTCTGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
 Qy 122 GCAAAATGCAAG 181
 Db 61 GCAAAATGCAAG 120
 Qy 182 GGTCAAGAGTGCAGAGAGATAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 241

D	b	121	GGTCAAGAGTACAGAGGAGTATGCAAGACCTGCTTACCAACATGTTCCTCGACAGAACGTGG	180
Q	y	242	GGGCCAGAGCGGGGCATCAGCCCTGACAGCCCATCGTACGTCTTGAGCTTGAGATCAC	301
D	b	181	GGGCCAGAGCGGGGCATCAGCCCTGACAGCCCATCGTACGTCTTGAGCTTGAGATCAC	240
Q	y	302	CAGCCAAACCTAGAGGCGCTGAGCCGCTTGCTGTCGCGCAGACGCCACAGATCTGAACTCAAG	361
D	b	241	CAGCCAAACCTAGAGGCGCTGAGCCGCTTGCTGTCGCGCAGACGCCACAGATCTGAACTCAAG	300
Q	y	362	GGCCCTGAGCAAGCTGAGCTGCTCATCCGGGAAACGCGACAGAGCTTCGCAAGACCTACAG	421
D	b	301	GGCCCTGAGCAAGCTGAGCTGCTCATCCGGGAAACGCGACAGAGCTTCGCAAGACCTACAG	360
Q	y	422	CGACAGTGGCAGCAGCTGCGACGAGAGCTCAACAAACCAACAGCCAGACATTGAGAA	481
D	b	361	CGACAGTGGCAGCAGCTGCGACGAGAGCTCAACAAACCAACAGCCAGACATTGAGAA	420
Q	y	482	GCTAAGAGCCAGTACCGAGCTCTGGGACCGGGACAGTSCCCAAAGCCCAAGTAAACA	541
D	b	421	GCTAAGAGCCAGTACCGAGCTCTGGGACCGGGACAGTSCCCAAAGCCCAAGTAAACA	480
Q	y	542	GGAGGCGCAGAAAGACAGAACCCGTGACAAAGGCCAAAGACAAAGTATGTGCCAGGCTGTG	601
D	b	481	GGAGGCGCAGAAAGACAGAACCCGTGACAAAGGCCAAAGACAAAGTATGTGCCAGGCTGTG	540
Q	y	602	GAAGCTCTTTGCTCAACAACAACCGGTATGTGTGGGCTGTGGGCTGGCAGCTACACCA	661
D	b	541	GAAGCTCTTTGCTCAACAACAACCGGTATGTGTGGGCTGTGGGCTGGCAGCTACACCA	600
Q	y	662	CCAGACCAACAACAACGCTCTCTGTGCTCCGGGCTGTGTGGGTCACTGACAGAACCTGCANGA	721
D	b	601	CCAGACCAACAACAACGCTCTCTGTGCTCCGGGCTGTGTGGGTCACTGACAGAACCTGCANGA	660
Q	y	722	GGAGTGGCTTGCATCTCGAAGAGATCTCTGACGAATACTGGAGATTACAGGCTGCT	781
D	b	661	GGAGTGGCTTGCATCTCGAAGAGATCTCTGACGAATACTGGAGATTACAGGCTGCT	720
Q	y	782	GCAGAGTAGGTTGTGTGGCCATTCAACCGGAGAGATGGCTGACGTGTGCTCCGCAATCAAGC	841
D	b	721	GCAGAGTAGGTTGTGTGGCCATTCAACCGGAGAGATGGCTGACGTGTGCTCCGCAATCAAGC	780
Q	y	842	TGAGGCTGAGTACCAAGGCTTCTGTGCAACAGTATGGGTCCGCACTGACGTGCCACCTGTG	901
D	b	781	TGAGGCTGAGTACCAAGGCTTCTGTGCAACAGTATGGGTCCGCACTGACGTGCCACCTGTG	840
Q	y	902	TGTCACTTTCATGATGACTCACTGCTTGAAGAGGGTCAACCGCTGAGGCTTGAGGAGCTTCCA	961
D	b	841	TGTCACTTTCATGATGACTCACTGCTTGAAGAGGGTCAACCGCTGAGGCTTGAGGAGCTTCCA	900
Q	y	962	GCTGACAGGCTGACTGTGTGAGAGCGGTGACAGACACGCTGACCTTCAGTGAACAGATGACT	1021
D	b	901	GCTGACAGGCTGACTGTGTGAGAGCGGTGACAGACACGCTGACCTTCAGTGAACAGATGACT	960
Q	y	1022	GAGTGTGGCCACCGAGATGTGTTCAGGCGGACAGAGATGTGTACGACGCTGCAACAGAA	1081
D	b	961	GAGTGTGGCCACCGAGATGTGTTCAGGCGGACAGAGATGTGTACGACGCTGCAACAGAA	1020
Q	y	1082	GCTCCGGAATGAAAGAGAGAACCCCAACCCCGGAGAGGGGTGACGCTGTGGGCAAGAG	1141
D	b	1021	GCTCCGGAATGAAAGAGAGAACCCCAACCCCGGAGAGGGGTGACGCTGTGGGCAAGAG	1080
Q	y	1142	GCAAGTGTGCAAGAGACCTGCGAGGGCTGACGGTATGCGCTGTGACAGCCAGGCCAACCT	1201
D	b	1081	GCAAGTGTGCAAGAGACCTGCGAGGGCTGACGGTATGCGCTGTGACAGCCAGGCCAACCT	1140
Q	y	1202	GCAAGGCCACAGAGAGATTGTGCAAGACCAAGCTGAGACACTTGTGGGCCCGCGAGACCCCC	1261
D	b	1141	GCAAGGCCACAGAGAGATTGTGCAAGACCAAGCTGAGACACTTGTGGGCCCGCGAGACCCCC	1200
Q	y	1262	GCGTGTGTGCTCTGCGAGAGTGAACCGCACTTCCACGTCTCTTGAGAGCAGAACCGAGA	1321
D	b	1201	GCGTGTGTGCTCTGCGAGAGTGAACCGCACTTCCACGTCTCTTGAGAGCAGAACCGAGA	1260

QY	1322	GGGGGAAAGGACACCCAGCGTGGAGATCCCTTAAGAGCAATCTCAGGAATCTTCCGCC	1381
Db	1261	GGGGGAAAGGACACCCAGCGTGGAGATCCCTTAAGAGCAATCTCAGGAATCTTCCGCC	1320
QY	1382	CAAGTTCTCG-----	1391
Db	1321	CAATTTCTCGTCCCTCCACCGGTGAGCTCATTTCCGAGGTGCAGAAAGCCCTGCATGA	1380
QY	1392	-----	1391
Db	1381	GCAGCTGTGGATCCACGGGGCCATCCCGAGGGCAGAGTGGCTGAGCTGCTGTGCACTC	1440
QY	1392	-----	1391
Db	1441	TGGGGACTCTCTGGTGGGGAGAGCCAGGGCAAGAGATAGTGTCTGTGGGTGCTGTG	1500
QY	1392	-----	1411
Db	1501	GGATGTCTGCCCCCGGCACTTCAATCAAGTCTTGGATTAATCTGTACCGACTGTGAAG	1560
QY	1412	GGAAAGGCTTTCCTAGCATTCCTTGGCTCATCGACCACTACTGAGACCAGAGCCCT	1471
Db	1561	GGAAAGGCTTTCCTAGCATTCCTTGGCTCATCGACCACTACTGAGACCAGAGCCCT	1520
QY	1472	CACCAAGAAGTGGTGTCTCTGCA CAGGGCTGTGCCCCAAGACAAAGTGGTGTCTGAA	1531
Db	1621	CACCAAGAAGTGGTGTCTCTGCA CAGGGCTGTGCCCCAAGACAAAGTGGTGTCTGAA	1680
QY	1532	CCATGAGAGCTGGTGTGTGGGTGAGAGATTTGGAACGGGGGAACCTTTGGCGAATGTTTCA	1591
Db	1681	CCATGAGAGCTGGTGTGTGGGTGAGAGATTTGGAACGGGGGAACCTTTGGCGAATGTTTCA	1740
QY	1592	CGGACGCGCTGCGAGCCGACAAACCCCTGTGGCGGTGAATCTTGTTCGAGAGACGCTCCC	1651
Db	1741	CGGACGCGCTGCGAGCCGACAAACCCCTGTGGCGGTGAATCTTGTTCGAGAGACGCTCCC	1800
QY	1652	ACCTGACCTCAAGGCCCAAGTTTCTACAGGAAGGAGATTCCTGAAGATACAGACCAACC	1711
Db	1801	ACCTGACCTCAAGGCCCAAGTTTCTACAGGAAGGAGATTCCTGAAGATACAGACCAACC	1860
QY	1712	CAACATCGTGGGTCTCATTTGGTGTCTSCACCCAGAGCACCCATCTACATCGTCAATGA	1771
Db	1861	CAACATCGTGGGTCTCATTTGGTGTCTSCACCCAGAGCACCCATCTACATCGTCAATGA	1920
QY	1772	GCTTGTGACGGGGGCGACCTTCTGACCTTCTCCGACGAGAGGGGCCCGCGTGGCGGT	1831
Db	1921	GCTTGTGACGGGGGCGACCTTCTGACCTTCTCCGACGAGAGGGGCCCGCGTGGCGGT	1980
QY	1832	GAAACATCTGTGTGAGATGGATGGGGGAGATGACAGCTGCTGGGATGGAGATACCTGGAGACAA	1891
Db	1981	GAAACATCTGTGTGAGATGGATGGGGGAGATGACAGCTGCTGGGATGGAGATACCTGGAGACAA	2040
QY	1892	GTGCTGCATCCACCGGAGACCTGGCTGTGCGAATCTGCTGTGTGACAGAGAAATGTCT	1951
Db	2041	GTGCTGCATCCACCGGAGACCTGGCTGTGCGAATCTGCTGTGTGACAGAGAAATGTCT	2100
QY	1952	GAAATCATGATACCTTTGGGATGTGCCAGAGGAAGCCGATGGGTCTATACAGCTTCAGG	2011
Db	2101	GAAATCATGATACCTTTGGGATGTGCCAGAGGAAGCCGATGGGTCTATACAGCTTCAGG	2160
QY	2012	GGGCGCTCAGACAAAGTCCCGGTGAAGTGAACCGCACCTGAAGGCCCTTAACTACAGGCGCTA	2071
Db	2161	GGGCGCTCAGACAAAGTCCCGGTGAAGTGAACCGCACCTGAAGGCCCTTAACTACAGGCGCTA	2220
QY	2072	CTCTCTCGAAAGGAGAGCTGTGAGCTTTGGCATTTGCTCTGGAGACCTTCAAGCTGTGG	2131
Db	2221	CTCTCTCGAAAGGAGAGCTGTGAGCTTTGGCATTTGCTCTGGAGAGACTTCAAGCTGTGG	2280
QY	2132	GGCGTCCCCCATCCCAACCTCAGCATCAGAGACACGGGAAGTTTGTGAGAAAGGGGGG	2191
Db	2281	GGCGTCCCCCATCCCAACCTCAGCATCAGAGACACGGGAAGTTTGTGAGAAAGGGGGG	2340

Qy	2192	CCGCTGCGCCGCCACAGCTGTGTCTCGATGCCGTGTTCAAGGCTCATATGAGACAGTGGT	2253
Db	2341	CCGCTGCGCCGCCACAGCTGTGTCTCGATGCCGTGTTCAAGGCTCATATGAGACAGTGGT	2400
Qy	2252	GGCCTATGAGCTGGGCGAGCGGCCAGCTTACAGCACATCTACACAGAGGCTGCAGAGCAT	2311
Db	2401	GGCCTATGAGCTGGGCGAGCGGCCAGCTTACAGCACATCTACACAGAGGCTGCAGAGCAT	2460
Qy	2312	CCGAAAGGGGAGTCGGTGAAGGCTGGAGACCCCTTCTCAAGCTGGTGGCTCTGCAGAGCCT	2371
Db	2461	CCGAAAGGGGAGTCGGTGAAGGCTGGAGACCCCTTCTCAAGCTGGTGGCTCTGCAGAGCCT	2520
Qy	2372	AGGTGACAGCTCCTCAGCGGCTTCAGACTCATATGCTGACAGCTCTTTCACAGTCTTGAGCTC	2431
Db	2521	AGGTGACAGCTCCTCAGCGGCTTCAGACTCATATGCTGACAGCTCTTTCACAGTCTTGAGCTC	2580
Qy	2432	CTGCGCACAGACATCCACACTGCGCGGACAGATGACAGCGCGGTGTCTCTCTGTGTCCCTGC	2491
Db	2581	CTGCGCACAGACATCCACACTGCGCGGACAGATGACAGCGCGGTGTCTCTCTGTGTCCCTGC	2640
Qy	2492	TGCTGCGCAGGGCTTCTCTTCCGGGCAAGAAACAATAAACAACATTGAGCCACTGAA	2548
Db	2641	TGCTGCGCAGGGCTTCTCTTCCGGGCAAGAAACAATAAACAACATTGAGCCACTGAA	2697

[illegible]

Query Match	58.0%;	Score 1552.2;	DB 10;	Length 2680;
Beet Local Similarity	79.1%;	Pred. No. 7.9e-228;		
Matches 1970;	Conservative 0;	Mismatches 303;	Indels 216;	Gaps 2;
OY	CGAAGCAGCACTAATGAGGCTTCTCTTCTGAGGTGACAGCCCCCAGGGCCACAGGGAGTCTCTG	119		
DB	48 CAGAGACACATTAATGAGGCTTCTCTTCAAGCTGTGACAGCCCCCAGGGCCACAGGGAGTCTG	107		
OY	120 CAGCAAAATGACAGAGGCGGAGCTTCGTCTACTGAGGGGATGAGAAATGAGATGAGCCAG	179		
DB	108 CAGCAGATGACGAGAAAGCTGAGCTGCGCTCTTAATGAGGGATGAGGAAGTGGATGGCCAG	167		
OY	180 CGGGTCAAGATGACAGGAGATATGACAGACTGTCTTACCACTATGTCCCTGACAGACATG	239		
DB	168 AGGGTCAAGATGACCGGGAAATATGACAGATTTGCTTACCACTATGTCTTGTGACAGACATG	227		
OY	240 GGGGGCCAGAGCCGGGGCCATCAGCCCTGACAGCCCCATCAGTCAAGTCTCTGGAGCTGAGATC	299		
DB	228 GGAGGCCAGAGCTGAGACAGCGGCCCCGACAGCCCTGTCAAGCTCTTGGGACAGAGATA	287		
OY	300 ACCAGCCAAATGAGGGCTGAGCGGCTGTGTCGGGACGACGCGAGAGATCTGAATCTCA	359		
DB	288 ACAAGCCAGACCGAGAACTTGAAGCCGGGTGTGCGGACGATGCAAGAAATCTGAATCTCG	347		
OY	360 GGGGCCCTGAGCAAGCTGAGCGCTGTCTCATCTCGGGAAAGCGACAGAGCTTCCGAAGCTCAC	419		
DB	348 GGGGCCCTTGAAGCAAACTGAAGGTGTCTATCTGGGAGCGGACCAAGCTGAGAAACGCTAC	407		
OY	420 AGCAGACAGTGGGACAGCTGACAGAGAGCTCACAAAGCCGACAGCAGAGGACATTTGAG	479		
DB	408 AACAGACAGTGGGACAGCTGACAGAGAGCTCACAAAGCCGACAGCAGAGGACATTTGAG	467		
OY	480 AAGCTGAAGGCCAGTACCCGAGCTCTGGCACCGGACAGTGCCCAAGCCAAAGCGCAAGTAC	539		
DB	468 AAGCTGAAGACTAGTACCGGACCTCTGTATGAGATGACCCAGGCCAGGCGCAAGTAC	527		
OY	540 CAGAGGCCAGCAAAAGACAGGACCGTGAACAGGCCCAAGGACATATGTGCGACAGCTG	599		
DB	528 CAGAGAACCGCAAAAGACAGGATCGAGACAGGCCCAAGGACATATGTGCGGAGCTGT	587		
OY	600 TGAAGGCTTTGTCTACCAACCGGTATGTGTGGGGCGTGGGGGCTGGCAGCTACAC	659		
DB	588 TGAAGGCTTTGTGCCACACACACCGCTGACGCTCTGGGATGAGGGCCGACAGCTGAC	647		
OY	660 CACCAAGCACACACACAGCTCTGCTGCGGGCTGTGCGGTGACATGACAGAGCTGAC	719		
DB	648 CACCAACACACACACCGGTTCATGCTGCTGCGCTGTGACATGACATGACAGATTTGAC	707		
OY	720 GAGGAGATGGCTTGATCTTGAAGAGATCTTGAGGAATATCTGAGATTTAGACGCTG	779		
DB	708 GAGGAGATGGCGGGCATTTCTGAAGACATCTTGAGGAATATCTGAGATTTAGACGCTG	767		
OY	780 GTGACAGATGAGGTGTGTGCTTACACCGGGAGATGTGCTGAGAGTGTGCTGCCGATCCAG	839		
DB	768 GTGACAGACGATGTGGACATTCATTCACCGTGAAGCTGAGAGTGTGCTGTGGAATTCAG	827		
OY	840 CCTGAGGCTGATGACCAAGGCTTCTGTGAGACAGTATGGGTCCGACCTGACGTCGCCACCC	899		
DB	828 CTTAGATTTGATGATCTTAAGGCTTCTGTGACAAATATGATTCACCTCTGATGTGCTACCT	887		
OY	900 TGTGTCAAGTTCAGATGATCACTGCTTGAAGAGGGTGAACCGCTGAGAGCTTGGGGAGTCT	959		
DB	888 TGTGTCACTTTTGAATGATGCTTCTTGAAGACGGGGAAACAGCTGAGACCGGGGAGCTG	947		
OY	960 CAGCTGAACGAGCTGATCTGTGAGAGGGGTGACACACGCTGACCTGACATGACAGATGAG	1019		
DB	948 CAGCTGAATGATTTGACCTTGAAGAGGGGTGACACACGCTGAATTTCTGTGACAGATGAA	1007		
OY	1020 CTGGCTGTGGCCACCGAGATGATGTTTCAAGGCGGACAGAGATGATTTACGACAGCTCAACAG	1079		
DB	1008 CTGGCTGTGGCCACCAAGAGATGCTGAGCCGCGACAGAGATGATTTACGACAGCTCAACGCT	1067		
OY	1080 GAGGCTCCGGAATGAAGAGAAACACCAACCCCGGAGCGGGTGCAGCTGCTGGGCAAG	1139		

1068 GAGCTCAAGTGAAGAGACCAACCCCGGGAGAGGGTGAAGCTTTCAGCAAG 1127
1140 AGGCAAGTGTGCAAGAGACCTGCAAGGGCTGCAAGTACGCTGTGCAAGCCAGGCGAAG 1199
1128 AGGCAATGTGTGCAAGAGCAATACAGGGGTGCAATAGCACTGTGCAAGCCAGGCAAG 1187
1200 CTGCAAGCCAGAGAGAGTGTGCAAGCAAGCTGAGAGCACTGGAGCCCGGCGAGGCC 1259
1188 CTTCAGGCCAGAGAGAGCTTGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1247
1260 CCGCTGTGTGTCTCTGCAAGAGTGAACCGCACTGCAAGTGTCTGTGCAAGAGAGAGAG 1319
1248 CAGAGCGGTGCGCTCTCTTCAAGAGTGAACCGCACTTCAAGCTCTCTCAAC-----GAGCGA 1301
1320 GAGGGGGAGAGAGAGACCCACGCTGAGAGATCTTAAAGAGCAATCTCAGAGATCTTCCG 1379
1302 GAGGGGGAGAGAGAGACCCACGCTGAGAGATCTTAAAGAGCAATCTTCCGAGATCTTCCG 1361
1380 CCGAAGTCTC----- 1390
1362 CCGAAGTCTCTATCCCTCCACCACTGCAAGCTGTGCGGAGGTGCAAGAGCTCTCTAT 1421
1391 ----- 1390
1422 GAGCAGCTGTGTACCAAGAGGGGCTATCCCTGGGCAAGAGGTAGCTGAGCTGTACAGCAC 1481
1391 ----- 1390
1482 ACTGGGAGACTTCTGTGTTCCGGAGAGCCAGGCAAGAGAGATGTACTGTGTGATG 1541
1391 -----GAACTGTACGAGCTGGA 1409
1542 TGGGATGGCAACCCCGACATTTCAATCATCACTGCTTAGATTAACCTGTACCACTGGA 1601
1410 GGGGAGAGCTTCTAGCATTTCTTGTCTATGCAACCACTTACTGAGCAACCCAGAGCC 1469
1602 GGGGAGAGCTTCTAGCATTTCTTGTCTATGCAACCACTTACTGAGCAACCCAGAGCC 1661
1470 CTCAACCAAGAGAGTGTGTGTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1529
1662 CTTCACAAAGAGAGTGTGTGTCTGTTTCAGGGCGGTGCGCAAGAGCAAGTGGGTCTG 1721
1530 AACCATGAGAGAGCTGT 1589
1722 AACCATGAGAGAGCTGT 1781
1782 AGTGGCGCTTGT 1841
1590 AGCGAGAGCTGT 1649
1650 CCACTGAGCTTCAAGGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1709
1842 CCACTGAGCTTCAAGGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1901
1710 CCAATGAGAGAGCTGT 1769
1902 CCAATGAGAGAGCTGT 1961
1770 GAGCTGT 1829
1962 GAGCTGT 2021
1830 GTGAAGACTGT 1889
2022 GTGAAGACTGT 2081
1890 AAGT 1949
2082 AAGT 2141
1950 CTGAAGT 2009

2142 CTGAAGT 2201
2010 GGGGGCTGT 2069
2202 AGCGGCTGT 2261
2070 TACTCTCCGAAAGAGAGCTGT 2129
2262 TACTCTCAGAGAGAGAGT 2221
2130 GGGGCTGT 2189
2322 GGGGCTGT 2381
2190 GGGGCTGT 2249
2382 CATGT 2441
2250 TGGGCTGT 2309
2442 TGGGCTGT 2501
2310 ATCCGAAAGCGCATGCTGT 2338
2502 ATCCGAAAGCGCATGCTGT 2530

RESULT 8
ACFTS140A 2715 bp ss-RNA linear VRL 27-APR-1993
LOCUS Fujinami sarcoma virus temperature sensitive (ts) p140 transforming
DEFINITION protein RNA, 3' end.
ACCESSION M14930
VERSION M14930.1 GI:209688
KEYWORDS transforming protein.
SOURCE Fujinami sarcoma virus
ORGANISM Fujinami sarcoma virus
REFERENCE 1 (bases 1 to 2715)
AUTHORS Chen, L.H., Hatada, B., Wheatley, W. and Lee, W.H.
TITLE Single amino acid substitution, from Glu1025 to Asp, of the fps oncogenic protein causes temperature sensitivity in transformation and kinase activity
JOURNAL Virology 155 (1), 106-119 (1986)
MEDLINE 87044080
PUBMED 2877522
COMMENT Original source text: Fujinami sarcoma integrated viral DNA, clone pFL-5 from ts FL-15.
FEATURES
source
CDS
1..2715
location/Qualifiers
/organism="Fujinami sarcoma virus"
/mol_type="genomic RNA"
/db_xref="taxon:11885"
<1..2697
/note="p140 transforming protein"
/codon_start=1
/protein_id="AAA42403.1"
/db_xref="GI:209688"
/translation="PDDVTNPKRVILGAPVLYLMDMAASGLRHPPOERTSTAA
AOTWRLTQASERHRLPHCSAASHQDSHAWGEPRLWCPRKSHSLRLQDSLRLE
LKKWMSQAKSDREYAGMLHMFQSLERQEGILHARDHSQIESWVLAQSTET
LSOTLRRAHEBLAAGLAKLILIRKQDLRVFSRQMOQLQSEYVMTLQOEVEKLKA
OYRSIVRSDTQAKRYQKASKDKBRKREKAYRSLSKLYALHNOVLAVOAALHNR
HHYQALPLTHSLKSLVSLQOEBMLVLKEIGYCSLTSLVQEBVLATHQBAALHND
PATERSSVQCHRYSVPEVPAVTFDSGLLEBASLEPGLQLNELTRESVQHSLEIB
EELLSRKAVSSKEQWELQVLELREBELALPGEVHLIGKRQGLRESAQQQLQGLVC
AQAQAQCRDMTLNKLAEISDEPPALPLQDRSAGCTDQERSGVTALKTKNHS
GIESPFSLPVPVLPVQKPLCOAWYHGAIPREVOELSLKSGDFLVRSEQKQK
VLSVIMDQBRHPTIILADNDLYLEDDGLPTLIDHLOSQRITRSGVILTRA
VAKDKRVNLHNEVULGERIGRGNPFGVSGRLRADNTPVAVVSCRETLPELKAFLQ
EARILKQCNHNPVRLIGVCTOKQPIYIWEVLQGGDFLFSLSKSPRLKMKULIOM
DNAAGMEYLSKHCHIRDLARNCLVTEKNTLKI SDPMSRQEDGQVASTGAMKOI
PVKMTAPBALNMGWSSBSDWVSFGILLWEAFSLGAVPYANLNSNOTREALIEQVRLIE

ORIGIN PPGCPEDVYRLMORCMEYDPHRRPSFGAHQDLAIRKMR"

Query Match 38.3%; Score 1024.8; DB 14; Length 2715;
Best Local Similarity 66.3%; Pred. No. 3.9e-147;
Matches 1650; Conservative 0; Mismatches 612; Indels 228; Gaps 3;

QY 64 ACAGCACTATGGGCTTCTCTCTGAGCTGTGACCCCGAGGGCCACGGGCTCTTGACG 123
DB 215 ACAGTCCATGGGCTTGGGCGGAGCTGTGTCCGAAAGGGGCACAGTGAAGCTCTGC 274
QY 124 AAATGAGAGAGGCGAGCTTCTCTCTACTGAGGGGCATGAGAAATGAGTGGCCAGCGG 183
DB 275 GGCTGAGAGACAGAGAGTTGCGCTCTCTGAGCTGTAGAGAAATGATGTACAGCTG 334
QY 184 TCAAGAGTACAGAGAGTATGACAGACTGTCTTACCAATGTCC-----CTGC 231
DB 335 CCMAAGCCACCGGAGTACGCGGGATGCTGCACACATGTTCTCTGAGCTGAGAAAC 394
QY 232 AGGACAGTGGGGCCAGAGCCGGGCCATCAGCCCTTACAGCCCATCACTGCTTGGG 291
DB 395 AGAGAGGCTTGGGACATCTCCGTGCCACCAAGCAGCAGATCGGGAGTCTGTGT 454
QY 292 CTGAGATCAACAGCAAACTGAGAGGCTGAGCCGCTTGTGCGGGCAGACGAGAGATC 351
DB 455 GGGTTTGGCAGACCACTGAGACACTGAGCCAGACGCTGCGGGCAGCGAGAGAGC 514
QY 352 TGAACCTCAGGGCCCTGAGCAAGCTGAGCTGTCTCCGGGAAACGGCAGAGCTTTCGA 411
DB 515 TGGCAGCGGGGGCGGTGGCCAGCTGAGTACTGATCCGTGACAAAGAGCTTTCGA 574
QY 412 AGACCTACAGCAGAGCTGTGACAGAGCTGACAGAGAGCTCACCAAGCCACAGCCAG 471
DB 575 AGGTCTTCAAGTGAAGCTGAGCAGAGCTCAAGCAGAGATACGCTGAGACCAACAGCAG 634
QY 472 ACATTGAGAGCTGAGAGAGCTGAGTACAGAGCTGTGCAAGGAGCATGACCAAGCCAG 531
DB 635 AAGTTGAGAACTGAGAGAGCAGTACCGAGCTGTGTGAGACAGACCCAGGCCAGAGC 694
QY 532 GCAAGTACAGAGAGCCAGCAAAAGACAGAGCCGTGACAAAGCCCAAGATATGTGC 591
DB 695 GCAAGTACAGAGAGCCAGCAAAAGACAGAGCCGAGAAAGGAGAAATAATGTGC 754
QY 592 GCAAGCTGTGAGAGCTTCTTGTCTACCAACCGCTATGTGTGAGGCTGTGCGGCTGCG 651
DB 755 GGAAGCTGTGAGAGCTTCTTGTCTTGTGCAACCAAGTACGTGTGTGAGGCGCTG 814
QY 652 AGCTACACCAAGCAGCAGCAGCTGCTGCGCGGCTGTGCGGCTGAGCTGAG 711
DB 815 CGTTGACACAGCCCATCTACCTACAGCGTGTGCGCACCGCTGACAGATCCCTTAC 874
QY 712 ACCTGACAGAGAGATGCTGTGATCTGAGAGAGATCTGACAGAAATCTGAGATTA 771
DB 875 GCTTGACAGAGAGATGATCTTGTGTTGAGAGATCTTGTGAGAGATCTGAGATTA 934
QY 772 GCAAGCTGTGAGAGTGAAGTGTGCGCAATTCACCGAGATGCTGAGCTGTGCGC 831
DB 935 CCAAGTGTGAGAGAGTGTGCGCAATTCACCGAGATGCTGAGCTGTGAG 994
QY 832 GCATCAGCTGAGAGTGTGAGTACCAAGGCTTCTGAGACAGTATGGGTCCGACCTGAG 891
DB 995 TGAATCAACCCCGCATGAGTACAGAGCTTGTGTGATGTCACAGATGAGCTTGAAG 1054
QY 892 TCCCAACCTGTGTGAGCTTGTGATGATCACTGCTTGAAGAGAGTGAACCGCTGAGAG 951
DB 1055 TACCAACCGGAGTGAAGCTTGTGATGAGAGCTGTGAGAGAGAGAGAGCTGAGAGC 1114
QY 952 GGAAGCTTCAAGCTGAAGAGCTGTGAGAGAGCTGAGACACAGCTGAGCTTCAAGTA 1011
DB 1115 GGAAGCTTCAAGCTGAAGAGCTGTGAGAGAGCTGAGAGAGCTGAGAGCTTCAAGTA 1174
QY 1012 CAGATAGCTGTGAGAGCTGTGAGAGAGTGTGAGAGAGAGAGAGAGTGTGAGAGAG 1071

DB 1175 AGAGAGAGCTTGTGAGCAGAGAAAGCAGTGAAGCAGAGAGAGAGAGAGAGCTGAGAGC 1234
QY 1072 TGCACAGAGAGCTCCGGAATGAAGAGAGAACACCAACCCCGGAGCGGCTGACAGTGC 1131
DB 1235 TCCAGTTGAGCTGCGGGCGAGAGCTGAGCAGTGAAGCTTGTGAGAGAGAGCTTGC 1294
QY 1132 TGGGCAAGAGAGAGTGTGAGAGAGAGCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGC 1191
DB 1295 TGGGCAAGC 1354
QY 1192 AGGCAAGCTGAGC 1251
DB 1355 AGGCAAGCTGAGC 1414
QY 1252 GCGAGCCCGCTGTGTGCTTCTGAGAGATGAGCCGACATTCAGTGTCTTGTGAGAGC 1311
DB 1415 AGAGAGCCCTGCGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1474
QY 1312 AGGAG 1371
DB 1475 AG 1528
QY 1372 TCTTCGCGCCCAAGTTCTCG-----CTGCGTGAAGACATCAAGAACCAATCTCGGAGC 1391
DB 1529 TCTTCAAGTCCAGAGTTCTGTGCTGCAACTCCGCTGCCATCCCGAGAGTGCAGAGAGC 1588
QY 1392 ----- 1391
DB 1589 CGCTGTCAGAGAGAGCTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1648
QY 1392 ----- 1391
DB 1649 TGAATACAGCGAGAGCTTCTGTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1708
QY 1392 -----AACTGTACC 1401
DB 1709 GGTGTGTGAGAGATGAGC 1768
QY 1402 GACTGAGC 1461
DB 1769 GGTGTGTGAGC 1828
QY 1462 AGCAGCCCTTCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1521
DB 1829 AGCAGCCCTTCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1588
QY 1522 GGTGTGTGAGC 1581
DB 1889 GGTGTGTGAGC 1948
QY 1582 AAGTGTTCAGCGAGC 1641
DB 1949 AGTGTGTGAGC 2008
QY 1642 AGAGCTTCCACCTGAGC 1701
DB 2009 AAGCTTCCACCTGAGC 2068
QY 1702 ACAGCAGC 1761
DB 2069 GCAACATCCCAATCTGTGCGGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 2128
QY 1762 TCGTCAATGAGC 1821
DB 2129 TCGTCAATGAGC 2188
QY 1822 GCTGTGAGC 1881
DB 2189 GCTGTGAGC 2248
QY 1882 TGAAGC 1941
DB 2249 TGAAGC 2308

QY	1942	AGAAATGCTCGAAGAAATCAGTGAACCTTTGGGAATGTCCTCCGAGAGGAACCCGATGGGGTCTATG	2001
Db	2309	AGAAACACCTCGTAAATACAGCACTTTGGGAATGTGCGCGCAGAGAGAGATGTCTATG	2368
QY	2002	CAGCCTCAGGGGGGCTCTCAGACAAATGCCCTCGTGAAGTGAACCGCACCTTAGAGCCCTTAAT	2061
Db	2369	CTTCAACGAGGGGGCATGAGACGATTCCTCGTGAATGATCTGCCCCCGAGGCTCTGAAT	2428
QY	2062	ACGGCCGCTACTCTCTCCGAAGCGACGTGTGGAAGCTTTGGCATCTTGTCTGGAGACT	2121
Db	2429	ACGCTGTGTAACAGCTCGGAGAGCAGCGTGTGAACCTTTGGGATCTCTGTGTGGAAAGCT	2488
QY	2122	TCAGCCTGGGGGCTCCCTCCATATCCAACTCCACGATTCAGACATTCAGACATCCGGAAGTTTGG	2181
Db	2489	TCAGCCTGGGGGCTCCCTCCATATCCAACTTCAGACATTCAGACATCCGGAAGTTTGG	2548
QY	2182	AGAAAGGGGGGCTCTGTGCTGCCCTCGACAGAGTGTGTCTGATGCGGTTCAGAGTCAATG	2241
Db	2549	AGCAGGGGGTGTGGGCTGGAGGCCCTCGAGAGTGTGTGTGATCCCGAGAGAGTGTATCAGCCTGATG	2608
QY	2242	AGCAGTGTGGGCTTATGAGCCTTGGGAGCGGCCCAAGCTTCAGACATTCAGACAGAGC	2301
Db	2609	AGCCTGTGTGGAGATCAGACCTTCAGAGCGCGCGAGCTTCGCGCGCTTCACAGAGACC	2668
QY	2302	TGCAGAGCATCCGAAAGCGGCATCGGTGAG	2331
Db	2669	TCATCGCATTCGCAACGAGCAGCGCTGAG	2698
RESULT	9		
ACF			
LOCUS		4788 bp ss-RNA	linear
DEFINITION		Fujinami sarcoma virus (unintegrated circular), complete genome.	VR1_27-APR-1993
ACCESSION		J02194 K01827 K01828	
VERSION		J02194.1 GI:209866	
KEYWORDS		C-myc proto-oncogene; complete genome; fps oncogene; polyprotein.	
SOURCE		Fujinami sarcoma virus	
ORGANISM		Fujinami sarcoma virus	
REFERENCE		1 (bases 1 to 4788)	
AUTHORS		Shibuya, M. and Hanafusa, H.	
TITLE		Nucleotide sequence of Fujinami sarcoma virus: evolutionary relationship of its transforming gene with transforming genes of other sarcoma viruses	
JOURNAL		Cell 30 (3), 787-795 (1982)	
MEDLINE		83050964	
PUBMED		6291784	
REFERENCE		2 (bases 1832 to 1881; 2851 to 2902)	
AUTHORS		Carlsberg, K., Chamberlin, M.E. and Beemon, K.	
TITLE		The avian sarcoma virus PRC1 lacks 1020 nucleotides of the fps transforming gene	
JOURNAL		Virology 135 (1), 157-167 (1984)	
MEDLINE		84425820	
PUBMED		6328746	
COMMENT		Original source text: FSV (Fujinami sarcoma virus), unintegrated circular DNA [1], [2].	
FEATURES		The sequence of 1182 amino acids was deduced from the Fujinami sarcoma virus (FSV) transforming protein p130, the product of gag-fps fused gene. p130 is highly homologous to the gag-fes sequence of feline sarcoma virus (see separate entries), and is homologous in the 280 residue carboxyl end (including the phosphoacceptor tyrosine residue) of Rous sarcoma virus p60 protein.	
SOURCE		Location/Qualifiers	
LTR		1..4788.	
repeat_region		/organism="Fujinami sarcoma virus"	
		/mol_type="genomic RNA"	
		/db_xref="taxon:11885"	
		1..347	
		/note="5' LTR"	
		247..267	
		/note="5' terminal repeat"	

Query Match	Best Local Similarity	Matches 1650; Conservative	38.3%;	Score 1024.8;	DB 14;	Length 4788;
QY	64	ACAGCACTATGGGCTTCTCTTGTGAGCTGTGCAGCCCCCAGGCGCCAGCGGGTCTGTGAGC	123			
Db	1692	ACAGTGCATGGGTTTGGGCCCGAGCTGTGTGTGCCGAAAGGGGCACTGAGCTGTGTC	1751			
QY	124	AAATGCAGAGGCCGAGCTTGTCTACTGTGAGGGGATGAGAAATGTGATGGCCAGCGGG	183			
Db	1752	GGCTGCAGGACAGCGAGTTGCGCTCTGTGAGGCTGATGAAGAAATGTGATTCACAGCGTG	1811			
QY	184	TCAAGAGTGCAGAGGATATGCAGAGATGTGTTACCACATGTGC-----CTGC	231			
Db	1812	CTAAGACCGACCGGAGATACGCGGGAGTGTGCACCAATGTTTCTTCACTGTGAGAAAC	1871			
QY	232	AGGACAGTGGGGGCGAGAGCCGGGCGCATTCAGACCTTCAGAGCCCGCATCATGTCCTGGG	291			
Db	1872	AGGAGGGCTTGGGACATCTCTGTCCACCCAGCACAGACGACAGATCGGGAGTCTGTGT	1931			
QY	292	CTGAGATCACAGCAAACTGAGGGCCCTGAGCCGCTTGTCTCGGCGACAGCGAGGATC	351			
Db	1932	GGGTTTGGCAAGCGACATGTAGACACTGAGCCAGACGCTGCGCGGCGACCGGAGGAGC	1991			
QY	352	TGAATCAGAGGCCCTTGTAGCAAGTGAAGCTGTCTCATTCGGGAAACGCGACGACTTTGCA	411			
Db	1992	TGGCAGCGGGGCGCTGGCCAAAGTGTGCACTATGATCTCGTGAACAGACAGCTTTGCA	2051			
QY	412	AGACCTTACAGAGAGTGGCAGAGCTGGAGCAGAGGCTCACCAAGGCCACAGCCAGG	471			
Db	2052	AGGCTTTCAGTGAACAATGGGACAGCTTCAGCCAGGAGTACGCTTGAACACACAGAGG	2111			
QY	472	ACATTGGAAGCTGGAAGGCCAGTACCGAGCTTGGCAGCGGACAGTGTCCCAAGC	531			
Db	2112	AAATTGGAAGCTGGAAGGCCAGTACCGAGCAGCTGTGTGTGACAGACCCAGGCGAAGC	2171			
QY	532	GCAAGTACAGAGAGCCAGCAAGACAGGACCTGTACCAAGGCCAAGAGCAAGTATGTGC	591			

Db 2172 GCAAGTACCAAGAGGCGGACGAGAACAGAGGCGGAAAGAGGAAAGGAAATGATCTGC 2231
Qy 592 GCAGCCTGTGGAAGCTTTTGTCTCAACCAACCGATATGTCTGTGGCGTGCAGGCTGC 651
Db 2232 GCAGCCTGTGGAAGCTTTATGCTTGTGCAACACAGTACGTGTGTGGCTGTGCAAGCGGCTG 2291
Qy 652 AGCTACCAACCAAGCAACCAACCAAGCTCTGTGCTCCGCGCTGTGCTGCTGCTGCAAG 711
Db 2292 CGTTGACCAACCAACCAACCAAGCTCTGTGCTCCGCGCTGTGCTGCTGCTGCAAC 2351
Qy 712 ACCTGACAGAGAGATGTGCTGTGCTCTGAAAGAGATCTGTGAGAAATACCTGAGATT 771
Db 2352 GCTGTGACAGAGAGATGTGCTGTGTTTGAAGAGATCTGTGGGAGTACTGTAGCATCA 2411
Qy 772 GCAGCCTGTGAGATGAGTGTGTGCTCAACCGGAGATGTGTGACGTGTGCC 831
Db 2412 CCAGTCTGTGTGAGAGAGATGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 2471
Qy 832 GCATCCAGCCTGTGAGTGTGATACCAAGCTTCTGTGCAAGTATGTGTGCTGCTGCA 891
Db 2472 TGATGACACCGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2531
Qy 892 TCCACCTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 951
Db 2532 TACACCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2591
Qy 952 GGGAGCTTCAAGCTGAAAGAGCTGATGTGAGAGCGTGTGAGCAACGCTGATGATG 1011
Db 2592 GGGAGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2651
Qy 1012 CAGATGAGT 1071
Db 2652 AGAGAGAGCTGT 2711
Qy 1072 TGCAACAGAGCTCCGGAATGAAAGAGAGAACCCACCCCGGAGCGGCTGTGAGTGC 1131
Db 2712 TGAGAGTGTGCTGT 2771
Qy 1132 TGGGCAAGAGGCAAGT 1191
Db 2772 TGGGCAAGGCGGAGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2831
Qy 1192 AGGCAAGCTGTGAGGCGGCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1251
Db 2832 AGGCAAGCTGTGAGGCGGCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2891
Qy 1252 GCGAGCGCGGCTGT 1311
Db 2892 AGAGAGCGGCTGT 2951
Qy 1312 AGAGAGGAGAGGAGGAG 1371
Db 2952 AGAGAGGAGAGGAG 3005
Qy 1372 TCTTCCGCGGCAAGT 1391
Db 3006 TCTTCCGCGGCAAGT 3065
Qy 1392 TCTTCCGCGGCAAGT 1391
Db 3066 CGCTGTGCAAGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3125
Qy 1392 TCTTCCGCGGCAAGT 1391
Db 3126 TGAATATACAGCGAGAGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3185
Qy 1392 TCTTCCGCGGCAAGT 1401
Db 3186 GCGT 3245
Qy 1402 GACTGGAAGGAGAGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1461

Db 3246 GCGTGAAG 3305
Qy 1462 AGAGAGCGGCTTCAACAG 1521
Db 3306 AGAGAGCGGCTTCAACAG 3365
Qy 1522 GCGTGTGAAG 1581
Db 3366 GCGTGTGAAG 3425
Qy 1582 AAGTGTTCAG 1641
Db 3426 AAGTGTTCAG 3485
Qy 1642 AGAGAGCGGCTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1701
Db 3486 AAGAGCGGCTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3545
Qy 1702 AGAGAGCGGCTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1761
Db 3546 AGAGAGCGGCTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3605
Qy 1762 TCGTCAATGAG 1821
Db 3606 TCGTCAATGAG 3665
Qy 1822 GCGTGTGAAG 1881
Db 3666 GCGTGTGAAG 3725
Qy 1882 TGAAG 1941
Db 3726 TGAAG 3785
Qy 1942 AGAGATGCTTGAAG 2001
Db 3786 AGAGATGCTTGAAG 3845
Qy 2002 CAGGCTGAG 2061
Db 3846 CAGGCTGAG 3905
Qy 2062 AGGAGCGGCTTCAACAG 2121
Db 3906 AGGAGCGGCTTCAACAG 3965
Qy 2122 TCAAGCTGTGAG 2181
Db 3966 TCAAGCTGTGAG 4025
Qy 2182 AGAGAGGAG 2241
Db 4026 AGAGAGGAG 4085
Qy 2242 AGAGAGCTGT 2301
Db 4086 AGAGAGCTGT 4145
Qy 2302 TGCAG 2331
Db 4146 TGCAG 4175

RESULT 10
AF033810
LOCUS AF033810 4788 bp RNA linear VRL 03-DEC-1998
DEFINITION Fujinami sarcoma virus, complete genome.
ACCESSION AF033810
VERSION AF033810.1 GI:2801466
KEYWORDS Fujinami sarcoma virus
SOURCE Fujinami sarcoma virus
ORGANISM
Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.

REFERENCE 1 (bases 1 to 4788)
AUTHORS Petropoulos, C.J.
TITLE Appendix 2: Retroviral taxonomy, protein structure, sequences, and genetic maps
JOURNAL (in) Coffin, J.M. (Ed.);
RETROVIRUSES: 757;
Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York,
NY, USA (1997)
REFERENCE 2 (bases 1 to 4788)
AUTHORS Chappey, C.
TITLE Direct Submision
JOURNAL Submitted (12-NOV-1997) NIH, NLM, Rockville Pike, Bethesda, MD
20894, USA
FEATURES
source .1. 4788
Location/Qualifiers
/organism="Pujiamt sarcoma virus"
/mol_type="genomic RNA"
/db_xref="taxon:11885"
22..4442
/product="p140 polyprotein"
22..102
380..3928
/codon_start=1
/product="p140 polyprotein"
/protein_id="AAC82565.1"
/db_xref="GI:2801467"
/translation="MEAVIKVTSACKTYCGKTSPEKRIQAMLSQLOKRLGILMSLD
LVSFGSDPITATLTORAMVYKSGELKTMGLVGLAKYAREQVTSBQAKMGLGSG
GRVSPPECTEKATERIKDGEITVQDITKAPREITAPKTVGSCITGHALICN
CATASAPPEPVYSGSLYPSLAVGEQGGDTPRGAEQPRAPRGVGTGLAPGALTDW
ARTREBLASTGPMVAMPVVIKTEGPAMPLEBKLJAGLAGAVAGSLSPAAVAVE
ALMSBLLPHDVTNPMRVILGPAPALMDAMASQGLRPOQHEITSAAAGTWRL
TQASRSRHLPHCSAASHODHSAMGFGELMCPKHTELLRLQDESLLELMKMM
SGAKSDREYAKMLHMFQSLKQSGHILRATDHSQJGESMTVIASTQETLSQTLR
RAHEBLAAGPLAKLSTILRDKQQLRVFSBQWQSLQETAMTQOEVETLAKQVSLV
RSTQAKRYQESKDKEREKAYRSLKYLALNLYVALVAALHHHRYGRA
LPTLSLSLYQEMVLVLEILGEYCSITSLYQEDVLAHQVAAVAMIDPATYS
SFVQCHRDSEVPVATPDSLEBAENLEPGLQNELTIEVHSLTISIEBLAS
KAVYSKEORVWEIOLVLEIRGEELASRGRVHLIGRQGLREAOQLOGLVCAQALQ
AQRDLANKLAELGSEPPALPLQSDROSASTDDBRSQVTRKTIKHIISGTSBR
FSLPPEVPLPEVQPLCOQAMTHGAIPESEVELKYSQDFLVRSGQKQEVYSLV
WDOQRHFIIOADNLRYLEDDELPTIPLIDHLOSQRITRKSQIVTRALVKQW
VNHEDVLLGERIGRNFGEVSGRLRQADTIPAVKSCRETLPELKAFLDEARLIK
OCHNPVIRLIGCTOKOPYIYVLEWOGDPLFSRSGRLMKKLIMENNAAG
MEYLSKHCHIRDLAARNCLVTEKNTLKISDFGMSROBEDGVYASTGSKQIPIVKTA
PEALNYGWYSSESDDWVSFGLLMEASLSLCAVPYANISNOCTRAIRISQRLRPBQCP
EDYRLMQRCMEYDPHRRSPFGAVHODLAIIRKRR"

mat_peptide
380..835
/product="p14 MA"
899..1087
/product="p10"
1085..1309
/product="CA"
1307..3925
/product="FBS"
4196..4442

mat_peptide
3'UTR

Query Match 38.3%; Score 1024.8; DB 14; Length 4788;
Best Local Similarity 66.3%; Pred. No. 4e-147;
Matches 1650; Conservative 0; Mismatches 612; Indels 228; Gaps 3;

QY 64 ACAGACTATGGGCTTCTTGTGAGCTGACGCCCCCAGGGCCACGCGGCTCTGTAGC 123
DB 1446 ACAGTGCATGGGCTTGGCGCGAGCTGTGTGCCGGAAGGGGACACTGAGCTGTGC 1505
QY 124 AAATGCAGAGCGCGAGCTTCTCTACTGAGGCGCATGAGAAATGGATGCCAGCGG 183
DB 1506 GGGTGCAGGACAGAGAGTTGGCTCTGAGCGTGAAGAGAGTGTATCAAGGCTG 1565
QY 184 TCAAGAGTGAACGAGGAGTATGACGAGCTTCAACCATGTCC-----CTGC 231
DB 1566 CTAAGAGCGACCGGAGGTACCGGGAGTGTGCAACCATGTCTCTCACTGGAAGAAC 1625

QY 232 AGACAGTGGGGGCCAGAGCCGGGCAATCATGCCCTGACAGCCCATCATGTCTCTGG 291
DB 1626 AGGAGGGCCTGGGACATCTCCCTGCCACCGACACAGAGCCAGATCGGGAGTCTGG 1685
QY 292 CTGAGATCAACAGCCAACTGAGGGCTTGAAGCTTGTCTGGCGACACAGAGAGATC 351
DB 1686 GGGTCTGGCAAGCACTGAGACTGAGACCTGAGCCGAGCGGCGGCGAGAGAC 1745
QY 352 TGAATCTAGGGGCCCTGAGCAAGCTGAGGCTGTCTATCCGGGAAAGGAGAGCTTCGA 411
DB 1746 TGGCAGCGGGGCGCTGGCCAGCTGAGCATCTGATCCGTGACAAAGACGCTTCGCA 1805
QY 412 AGACCTACAGAGAGAGTGGCAGAGCTGACAGAGAGCTCAACAGACCCAGCCAG 471
DB 1806 AGGTCTTCACTGATGAGCAGTGGCAGAGCTCAGCGCAGAGATACCGCTGGAACACAGCAG 1865
QY 472 ACATTGAGAGCTGAAGAGCCAGTACCGAGCTCTGGCACGGACAGTCCCAAGCCAGC 531
DB 1866 AAGTTGAGAGAGCTGAAGGCAAGTACCGCAGCTGTGCTGTGACAGCACCCAGGCCAAC 1925
QY 532 GGAAGTACAGAGAGCCAGCAAGAACAGAGACCGTGAAGAGCCAGGCAAGATATGTC 591
DB 1926 GGAAGTACAGAGAGCCAGCAAGAACAGAGACCGGAAAGGCGAAGAAAGTACGTGC 1985
QY 592 GCAGCCTGTGAAGCTCTTGTCTCACCAACACCGCTATGTGTGGCTGTGGCGCTGC 651
DB 1986 GGAAGCTGTGAAGCTCTATGCTTGTGACCAACAGTATGTGTGTGTGAGAGCGGCTG 2045
QY 652 AGCTACACCAACAGACACCAACAGCTCTGTGCGCGCTGTGCGGTGATCTGCAAG 711
DB 2046 CTTTGACACCAACACATCACTACAGCGTCTGTGCGCCACCTGTGACAGATCCCTTACA 2105
QY 712 AACTGACAGAGAGATGCTTCATCTCTGAAGAGATCTGTGAGAAATCTGTGAGATTA 771
DB 2106 GCTTGACAGAGAGATGCTCTTGTGTAAGAGATCTTGGGAGATCTGTAGATCA 2165
QY 772 GCAGCCTGTGAGAGTGAAGTGTGTGGCCATTACCGGAGATGTGCTGACGTGCTGCC 831
DB 2166 CCAAGTGTGTGACAGAGAGCGTGTGCGATCAACAGAAAGTGTGCCAGCGCGTGAAG 2225
QY 832 GCATTCAGCTTGAAGCTGATACCAAGGCTTCTGTGCAAGTATGGTCCGACCTGACG 891
DB 2226 TGAATCAACCCGCCCACTGAGTACAGACGCTTGTGTCAGTCCACAGATACGATCTGAGG 2285
QY 892 TCCACCCCTGTGACGCTTGAATGAGTCACTGCTTGAAGAGGAGTGAACCGCTGAGCCTG 951
DB 2286 TACCAACCGGACGATGACCTTGAATGAGAGCTTGTGAGAGAGGAGAACTGTGAGCCAG 2345
QY 952 GGGAGCTCCAGCTGAACGAGCTGATGTGAAGAGCGTGAAGCAACAGCTGATCTCACTGA 1011
DB 2346 GGGAGCTGAAGCTGAATGAGTGAACATGAGAGCGGTGAACACATCCCTGATCATCTTGG 2405
QY 1012 CAGATGAGCTGGCTGTGGCCACCGAGATGTGTTACGGGGGAGAGATGTATCGCAGC 1071
DB 2406 AGAGAGAGCTGTGTGCAAGAGAGAGAGTGAAGAGAGAGAGCGAGGAGTGTGAGAG 2465
QY 1072 TGAACAGAGAGCTCCGGAATGAAGAGAGAGAACCCACCCCGGAGCGGGTCACTGC 1131
DB 2466 TGAAGTTAGAGCTGCGGGGCGAGAGACTGTGCACTTAGCCCTTGGGAGCGGGTCACTTGC 2525
QY 1132 TGGGCAAGAGAGCAAGTGTGCAAGAGCACTGACGAGGGCTGTGAGTATGCGTGTGAGCC 1191
DB 2526 TGGGCAAGGCGGAGGGGCTGCGGGAGGCCAGAGAGAGCTGCAAGGCTTGTGTGCTC 2585
QY 1192 AGGCAAGCTGACGAGCCACAGAGAGTGTGTCAGACCAAGCTGAGACACTGTGGCCCG 1251
DB 2586 AGGCAAGCTGACGAGCAACGCGGATATCTGGCCAAAGCTGTGAGAGCTGTGGCTGTG 2645
QY 1252 GCGAGCGCCCGCTGTGCTGCTCTGCAAGATGACCGGCACTCAGTGTGCTCGGAGC 1311
DB 2646 AGGAGCCCTCCCTCGCTTGTGCTTGTGCAAGAGAGCGGAGTGAAGCGCTTCAAGGATC 2705

Db 301 GGGTCTATGACGCTCAGGGGGGCTCAGACAAAGTCCCCGTAAGTAGACCGACCTGAGG 360
Qy 2053 CCGTTAACTACGGCCCTACTCTCCGAAAGCGACGTGTGACCTTTGGCATTTGCTCT 2112
Db 361 CCTTAACTACGGCCCTACTCTCCGAAAGCGACGTGTGACCTTTGGCATTTGCTCT 420
Qy 2113 GGGAGACCTTCAAGCTGGGGGCTCCCTCATCCCAACTCAGCAATCAGACAGACGGG 2172
Db 421 GGGAGACCTTCAAGCTGGGGGCTCCCTCATCCCAACTCAGCAATCAGACAGACGGG 480
Qy 2173 AGTTTGTGAGAAAGGGGGGCGGTGTGCTCCCAAGAGCTGTCTGATGCGGTGTCA 2232
Db 481 AGTTTGTGAGAAAGGGGGGCGGTGTGCTCCCAAGAGCTGTCTGATGCGGTGTCA 540
Qy 2233 GGGTCAATGAGAGAGTGTGGGCTATGAGCTGTGGGAGCGGCGCCAGCTTCAGACCATCT 2292
Db 541 GGGTCAATGAGAGAGTGTGGGCTATGAGCTGTGGGAGCGGCGCCAGCTTCAGACCATCT 600
Qy 2293 ACCAGAGCTGACAGACATCCGAAAGCGGATCGTGTGAGGCTGGGACCCCTTCTGAGC 2352
Db 601 ACCAGAGCTGACAGACATCCGAAAGCGGATCGTGTGAGGCTGGGACCCCTTCTGAGC 660
Qy 2353 TGGTGCCTCTGACAGGCTTATGATGAGCTCTCAGCGGCTCCAGCTCATATGCTGACAGC 2412
Db 661 TGGTGCCTCTGACAGGCTTATGATGAGCTCTCAGCGGCTCCAGCTCATATGCTGACAGC 720
Qy 2413 TCTTCAAGTCCGAGACTCCTGACACAGATCACAATGCGGGAGGAATGACGGCGCT 2472
Db 721 TCTTCAAGTCCGAGACTCCTGACACAGATCACAATGCGGGAGGAATGACGGCGCT 780
Qy 2473 GTCTCTCTGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2532
Db 781 GTCTCTCTGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy 2533 ACTTGTGCCCACTGAA 2548
Db 841 ACTTGTGCCCACTGAA 856

RESULT 12
M22820
LOCUS
DEFINITION
M22820 4901 bp DNA linear UNA 04-AUG-1993
Figure 4. Nucleotide sequence of the proviral DNA of the variant
clone of ST-FesV.
M22820.1 GI:340749
ACCESSION
M22820.1
KEYWORDS
unclassified
SOURCE
unclassified
ORGANISM
unclassified
REFERENCE
1 (bases 1 to 4901)
AUTHORS
Van Groningen,J.J., Van den Ouweland,A.M., Verbeek,J.S., van der
Kemp,A.W., Bloemers,H.P. and Van de Ven,W.J.
TITLE
Structural analysis of a variant clone of Snyder-Theilen feline
sarcoma virus
JOURNAL
Virus Res. 8 (4), 349-361 (1987)
MEDLINE
88129048
PUBMED
2829460
FEATURES
Location/Qualifiers
1..4901
/organism="unclassified"
/mol_type="unassigned DNA"
/db_xref="taxon:12908"

ORIGIN
Query Match 32.0%; Score 854.4; DB 13; Length 4901;
Best Local Similarity 77.7%; Pred. No. 4,7e-121;
Matches 1159; Conservative 0; Mismatches 121; Indels 212; Gaps 2;

Qy 1104 ACCCAACCCCGGAGCGGGTGTGACAGTCTGTGGCAAGAGGCAAGTGTCTGCAAGAGCACTG 1163
Db 1874 ACCCAACCCCGGAGCGGGTGTGACAGTCTGTGGCAAGAGGCAAGTGTGTGCAAGAGCGCTG 1933

Qy 1164 CAGGGGCTGACAGTAGCGTGTGACCCAGGCGCAAGCTGACAGGCCAGGAGATTGCTG 1223
Db 1934 CAGGGGCTGACAGTAGCGTGTGACCCAGGCGCAAGCTGACAGGCCAGGAGATTGCTG 1993
Qy 1224 CAGACCAAGCTGAGAGCCTGGGCCCCGGGAGAGCCCGGCTGTGCTCTGTCAGAGAT 1283
Db 1994 CAGGCGCAAGCTGAGAGCCTGGGCCCCGGGAGAGCCCGGCTGTGCTCTGTCAGAGAT 2053
Qy 1284 GACCGGCACTGACAGTGTGCTCGGAGCAGAGCGAGAGGGGGAGAGCAACCAAGCTG 1343
Db 2054 GACCGGCACTGACAGTGTGCTCGGAGCAGAGCGAGAGGGGGAGAGCAACCAAGCTG 2113
Qy 1344 GAGATCTTTAAGAGCACAATCTCAGGAATCTTCGCGCCCAAGTTCG----- 1391
Db 2114 GAGATCTTTAAGAGCACAATCTCAGGAATCTTCGCGCCCAAGTTCGCTCCCTCCAGCC 2173
Qy 1392 ----- 1391
Db 2174 CTGACGCTGTATCAGAGGTGAGAAAGCCCTGACAGAGAGCTGTGTATCAACGAGGCGC 2233
Qy 1392 ----- 1391
Db 2234 CTCCACGGGACAGAGGTGAGCTGTGACGCACTCTGGGAGCTTTGCTGGCGGAG 2293
Qy 1392 ----- 1391
Db 2294 AGCAAGGCGAGAGGATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2353
Qy 1392 -----AACTGTATCCGACTGTGAAAGGGGAAAGGCTTTCTAGCATTTCT 1433
Db 2354 ATCATTCAGTCCGCTGACAACTCTACCGATGTGAAAGGAGGTGTCTTGTGAGCATCTCC 2413
Qy 1434 TTGCTCATGACACACTTACTGAGACCCAGAGCCCTTACCAAGAGAGTGTGTGTGTC 1493
Db 2414 TTGCTCATGACACACTTACTGAGACCCAGAGCCCTTACCAAGAGAGTGTGTGTGTC 2473
Qy 1494 CTGACAGGCTGTGTGCCAAGAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1553
Db 2474 CTCAACAGGCTGTGTGCCAAGAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2533
Qy 1554 GAGCAGATTGACAGGGGGAATTTGGCGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1613
Db 2534 GAGCAGATTGACAGGGGGAATTTGGCGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2593
Qy 1614 ACCCTGTGGGCGGTGAGCTTGTGTGAGAGCGCTCCACTGTGACTCAAGGCGCAAGTTT 1673
Db 2594 ACTTATGTGCGGTGAAATTTGTGTGCGAGACATCTCCACTGTGACTCAAGGCGCAAGTTT 2653
Qy 1674 CTACAGAGAGCAGAGATCTGTGAGAGCAAGTACAGCAACCCCAACATCGTGTCTCATTTGT 1733
Db 2654 CTTCAGAGAGAGCAAGATCTGTGAGAGCAAGTACAGCAACCCCAACATCGTGTCTCATTTGT 2713
Qy 1734 GTCTGACCCAGAGAGAGCGCCATCTACATGCTGTATGAGCTTGTGTGAGGGGGCGACTTC 1793
Db 2714 GTCTGACCCAGAGAGAGCGCCATCTACATGCTGTATGAGCTTGTGTGAGGGGGCGACTTC 2773
Qy 1794 CTGACCTTCTCTCGAGAGGGGGGCGGCTGTGAGGGGGTGAAGCTGTGTGAGATGATG 1853
Db 2774 CTGACCTTCTCTCGAGAGGGGGGCGGCTGTGAGGGGGTGAAGCTGTGTGAGATGATG 2833
Qy 1854 GGGGATGACAGCTGTGTGAGAGTACTGTGAGAGCAAGTGTGTGTGTGTGTGTGTGTGTGT 1913
Db 2834 GGGGATGACAGCTGTGTGAGAGTACTGTGAGAGCAAGTGTGTGTGTGTGTGTGTGTGTGT 2893
Qy 1914 GCTGTCTGGAATCTGCTGT 1973
Db 2894 GCTGTCTGGAATCTGCTGT 2953
Qy 1974 TCCCGAGAGAGAGCGGATGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2033
Db 2954 TCCCGAGAGAGAGCGGATGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3013
Qy 2034 AAGTGACCGCACTGTAGGCGCTTAACTACGGCGGCTACTCTCTCGAAAGCGAGGTGTGG 2093

QY 1674 CTACAGAAAGCGAGATCTGAAGACAGTACAGCCACCCCAACATCGTGCCTCATTTGAT 1733
 DB 1669 CTTACAGAAAGCGAGATCTGAAGACAGTACAGCCACCCCAACATCGTGCCTCATTTGAT 1728
 QY 1734 GTCGTGACCCAGAGAGAGCCCATCTACATGTCATGAGAGCTTGTGACAGGGGGGAGACTTC 1793
 DB 1729 GTCGTGACCCAGAGAGAGCCCATCTACATGTCATGAGAGCTTGTGACAGGGGGGAGACTTC 1788
 QY 1794 CTGACCTTCTCCGACAGAGAGGGGGCCCGCTCCGAGTGAAGCTGCTGTCAGATGATG 1853
 DB 1789 CTGACCTTCTCCGAGAGAGAGGGGGCCCGCTCCGAGTGAAGAGCTGCTGTCAGATGATG 1848
 QY 1854 GGGAGATGACAGCTCTGAGATGAGATGCTGAGAGAGAGTGTCTGATCCACCGAGACTTC 1913
 DB 1849 GGGAGAGAGGGGGGGGGGGAGAGTACCTGAGAGAGAGAGTGTCTGATCCACCGAGACTTC 1908
 QY 1914 GCTGCTCGAAGCTGCTGCTGAGAGAGAGAGAGAGTGTCTGAGAGAGTGTCTGAGAGATG 1973
 DB 1909 GCTGCTCGAAGCTGCTGCTGAGAGAGAGAGAGAGTGTCTGAGAGAGTGTCTGAGAGATG 1968
 QY 1974 TCCCGAG 2033
 DB 1969 TCCCGAG 2028
 QY 2034 AAGTGAACGACAGCTGAG 2093
 DB 2029 AAGTGAACGACAG 2088
 QY 2094 AGCTTTGGACATCTTGTCTGAG 2153
 DB 2089 AGCTTTGGACATCTTGTCTGAG 2148
 QY 2154 AGCAATCAG 2213
 DB 2149 AGCAATCAG 2208
 QY 2214 TGTCTGAG 2273
 DB 2209 TGTCTGAG 2268
 QY 2274 CCGAGCTTACAGACATCTACAG 2333
 DB 2269 CCGAGCTTACAG 2328
 QY 2334 TGGAG 2385
 DB 2329 -CGAGCTTGTCTCAAG 2379

RESULT 14
 FCGGAGAG 2946 bp ss-RNA linear VRL 02-AUG-1993
 LOCUS Feline sarcoma virus (gardner-arnstein) transforming gene.
 DEFINITION J02087.1 GI:323872
 ACCESSION J02087
 VERSION J02087.1
 KEYWORDS c-myc proto-oncogene; fcs oncogene; polyprotein.
 SOURCE Gardner-Arnstein feline leukemia oncovirus B
 ORGANISM Gardner-Arnstein feline leukemia oncovirus B
 RETROVIRUSES; 1-Mammalian type C virus group.
 1 (bases 1 to 2946)
 Hamp, A., Laprevotte, I., Galibert, F., Fedele, L.A. and Sherr, C.J.
 Nucleotide sequences of feline retroviral oncogenes (v-fes) provide
 evidence for a family of tyrosine-specific protein kinase genes
 Cell 30 (3), 775-785 (1982)
 JOURNAL MEDLINE 83050963
 PUBMED 6183005
 COMMENT Original source text: fsv (feline sarcoma virus) gardner-arnstein
 strain.
 the transforming polyprotein of the gardner-arnstein strain of
 feline sarcoma virus (fsv) include the transforming gene (fes)
 product with tyrosine specific protein kinase activity. the encoded
 protein sequence was compared with the gene products of

anyder-theilen strain (see fcsv onc), fcs gene of avian fujinami
 sarcoma virus (see fsv), and src gene of avian sarcoma virus. the
 last 24 bases represent the sequence of feline leukemia virus.

FEATURES
 source
 1..2946
 /organism="Gardner-Arnstein feline leukemia oncovirus B"
 /mol_type="genomic RNA"
 /db_xref="taxon:11824"
 1..2874
 /note="gag polyprotein"
 /codon_start=1
 /protein_id="AAA43041.1"
 /db_xref="GI:323873"
 /translation="MGQITTPPLSTLDMSEVRAHNOGVEKKWITLCEAWV
 MNVGMPEGTESLNIGVEKTIAPAGPGYGDVQVYITTRSLATDPSWVRLP
 PKPTPTSLPQSLPPSPAPLSTSLYPLVLPKSPKPPVLPDPSRLIDLITREPPY
 PGHGPPSPGPTPTASPLASLIRREKREPAESQPLREBPNNRPQVTPWSADLY
 NMKSNPPSPDPPALTNLISILVTHQPTWDDCOQLQALLGSEERQVLEARKQV
 PSEDRPTOLPVVIDETPLTPNMDPAPACBHLRLRLQLLGLRGAARPNVLA
 QVKOVVQGEETPAARADGTWGFSSSELCS PQGHAEQOMQAELELLEGMKMAOR
 VSKDEYVGLIHHMSLIDGGGRTGTPYSPISQSWAEITSTQEGSLRLRQHAEDNSG
 PLSKGLIRERQQLRTYSBEMQLOELTTHNODLEKLSOYRALRDSQARRK
 YQASKDXDRDAKLEQLGPGEPVPLVLDNRHSTSSBQREBGRPTTLTSLHI
 SGIFRPKPSLPEPLQLVPEVQPLRHOULMHGALPRAEVALITHSQFLVESQKQ
 EYLVSLMDGQPRHPIIOSADNLVPEGDFASIPLDVHLRSQQLPTKSGIYVLR
 AVPKDMLNHEDLVIGROI GRNGEVEVSGRLADNLTVAVSGREPTLPDIKAFI
 QBAKILKQYSHNIVRLIGVCRQXQPIYVMEILVGGGDLPLRTFGAELRMKTLLOM
 VGDAAAGMEYLESKCIHRDLAARNCITLREKVLITSDGMRBAADGIVASGGLRO
 VPKNTAFBALMYGSSSDWSYSGILMETFSLGASGTPYPLNSQITREFEKGRL
 PCPELCPPAVFLMEQCMAYEPGPPSPALVQELQSLIKRRR"
 4..381
 /product="protein p15"
 382..591
 /product="protein p12"
 592..1044
 /product="protein p30"
 1045..2871
 /product="v-fes transforming protein"
 ORIGIN about 200bp upstream from a bglII site.

Query Match 31.1%; Score 831.2; DB 14; Length 2946;
 Best Local Similarity 82.8%; Pred. No. 1.6e-117;
 Matches 987; Conservative 0; Mismatches 198; Indels 7; Gaps 3;

QY 1200 CTGCAAGGCGCCAGAGAGAGTGTGACAGACCAAGCTGAGACACTGAGG---CCCGGCGAG 1256
 DB 1738 CTGCAAGCTGTACAGAGAGTGAAGAGCCCTGACAGACAGTGTGTACCAAGGAGGCC 1797
 QY 1257 CCGCGGCTGTGCTGCTCTGAGAGATGAACGCACTTCACGTGCTCTGAGAGAGAG 1316
 DB 1798 CTCCCAAGGCGAGAGAGTGTGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1857
 QY 1317 CGAAGAGGCGGAG 1376
 DB 1858 AGCAGAGGCGAG 1917
 QY 1377 CGGCCCAAGTTTC---GAACCTGTACGAGCTGGAAGAGGAGAGAGAGAGAGAGAGAGAGAG 1433
 DB 1918 ATCATCAGTCCGCTGACACACTCTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1977
 QY 1434 TTGCTCATGACACACTCTGACAG 1493
 DB 1978 TTGCTGTGTGACACACTGTGAG 2037
 QY 1494 CTGACAGAGGCTGTGCTGAG 1553
 DB 2038 CTCAACAGAGGCTGTGCTGAG 2097
 QY 1554 GAGCAGATGAG 1613
 DB 2098 GAGCAGATGAG 2157
 QY 1614 ACCCTGTGCGAGTGAAGCTTGTGAG 1673

DB 2158 ACCTGTGTCCTGTAATCTTCTCCGAGACACTCCCACTGACATCAAGGCCAAGTTT 2217
 QY 1674 CTACAGAGAACGAGATCTTGAAGACGTAACGACACCCCAACATGCTGCTCATTTGTT 1733
 DB 2218 CTTACAGAGAACAAAGATCTCGAAGAGATACAGCCACCCCAACATCGTGGCTCATCGGC 2277
 QY 1734 GTCCTGACCCAGAGAGAGCCCATCTACATGCTATGAGAGCTTGTGAGAGGGGGCGACTTC 1793
 DB 2278 GTCCTGACCCAGAGAGAGCCCATCTACATGCTATGAGAGCTTGTGAGAGGGGGCGACTTC 2337
 QY 1794 CTGACCTTCTCCGACGAGAGGGGGCCGCTCGGGGTGAAGACTGTGCTGACATGTGTG 1853
 DB 2338 CTGACCTTCTCGTAGAGACGAGAGGAGCCCGCTCGGATGAAGACGCTGCTGACGATGTG 2397
 QY 1854 GGGGATGACAGCTGCTGAGATGTAAGTCTGAGAGAGAGTGTGATCCACCGGAGCTTG 1913
 DB 2398 GGGGATGACGAGGCGCGGGGATGAGTACTGAGAGAGCAAGTGTGATCCACCGGAGCTTG 2457
 QY 1914 GCTGCTCGGACCTGCTGTGAGAGAGAAATGTCTGAAGATCAATGATCTTGGGATG 1973
 DB 2458 GCTGCTCGGACCTGCTGTGAGAGAGAAATGTCTGAAGATCAATGATCTTGGGATG 2517
 QY 1974 TCCCGAGAGAGAGCCGATGGGGGTCTATGACGCTCAAGGGGGCTCAAGACAAAGTCCCGTG 2033
 DB 2518 TCCCGAGAGAGAGCCGATGGGGGTCTATGACGCGGCTCAAGGGGGCTCAAGACAAAGTCCCGTG 2577
 QY 2034 AAGTGAACGCGACCTGAGGCGCTTAACTACGAGCGGCTACTCTCGAAGACGAGTGTG 2093
 DB 2578 AAGTGAACGCGACCTGAGGCGCTTAACTACGAGCGGCTACTCTCGAAGACGAGTGTG 2637
 QY 2094 AGCTTTGGCATCTTGTCTTGGAGACCTTCAAGCTTGGGGGCTTCCCTATGCCAACCTTC 2153
 DB 2638 AGCTTTGGCATCTTGTCTTGGAGACCTTCAAGCTTGGGGGCTTCCCTATGCCAACCTTC 2697
 QY 2154 AGCAATCAGACAGACGAGAGTTTGTGAGAGAGGGGGGCGCTGCGCCCTGCGCAGAGCTG 2213
 DB 2698 AGCAATCAGACAGACGAGAGTTTGTGAGAGAGGGGGGCGCTGCGCCCTGCGCAGAGCTG 2757
 QY 2214 TGTCTGATGTCCTGTTTCAAGCTCATGAGAGTGTGGGCTATGAGCTTGGGACGCG 2273
 DB 2758 TGTCTGATGTCCTGTTTCAAGCTCATGAGAGTGTGGGCTATGAGCTTGGGACGCG 2817
 QY 2274 CCGAGCTTCAAGACCATCTTACGAGAGCTGCAAGACATCCGAAAGGGGACGCTGAGGC 2333
 DB 2818 CCGAGCTTCAAGACCATCTTACGAGAGCTGCAAGACATCCGAAAGGGGACGCTGAGGC 2877
 QY 2334 TGGGACCCCTTGTCAAGCTGATGGGCTTGCAGAGCCCTGAGAGTGCAGAGCTC 2385
 DB 2878 -CGGCCCTGCTTCTCAAGCCGCGGCTCGACAGGCAAGTGTACTCTC 2928

RESULT 15
 LOCUS BC073445 4788 bp mRNA linear VRT 03-AUG-2004
 DEFINITION Xenopus laevis MG80946 protein, mRNA (cDNA clone MGC:80946
 IMAGE:5516233), complete cds.
 ACCESSION BC073445
 VERSION BC073445
 KEYWORDS MGC.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus; Xenopus.
 1 (bases 1 to 4788)
 Kleier, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.
 and Richardson, P.
 Genetic and genomic tools for Xenopus research: The NIH Xenopus
 Initiative
 JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
 PUBMED 12454917
 REFERENCE 2 (bases 1 to 4788)

AUTHORS Strausberg, R.L., Peingold, B.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shemmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, D., Hsieh, F.,
 Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carinci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mulhally, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahy, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Sanchez, A., Whiting, M., Madan, A., Touchman, J.W., Green, E.D.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalske, U., Smallus, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 3 (bases 1 to 4788)
 Klein, S. and Gerhardt, D.S.
 Direct Submission
 Submitted (14-JUN-2004) National Institutes of Health, Xenopus Gene
 Collection (XGC), National Institute of Child Health and Human
 Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
 20892-7510, USA
 NIH-MGC Project
 Contact: XGC help desk
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Dr. Igor David
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@gscc.bc.ca
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
 Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
 Featherstone, Malachi Griffith, Ohi Griffith, Ran Guin, Nancy Liao,
 Kim Macdonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
 Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lisa Prabhu,
 Parvaneh Seedi, JR Santos, Angeliqne Scherch, Ursula Skalske,
 Duane Smallus, Jeff Stolt, Miranda Tsai, George Yang, Jacque
 Schein, Asim Siddiqui, Rob Holt, Marco Marra.

FEATURES
 source
 1..4788
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="MGC:80946 IMAGE:5516233"
 /tissue_type="Embryo, stage 31/2, Xenopus"
 /clone_lib="NICHD_XGC_Embd4"
 /lab_host="PH10B"
 /note="Vector: pCMV-SPORT6"
 1..4788
 /gene="MGC80946"
 /db_xref="LOCUSID:444292"
 399..2867
 /gene="MGC80946"
 /codon_start=1
 /product="MGC80946 protein"
 /protein_id="AA073445.1"
 /db_xref="GI:49116711"
 /db_xref="LOCUSID:444292"

CDs
 gene
 title
 journal
 pubmed
 refseq
 2 (bases 1 to 4788)


```
Db      2256 GTCAAACTTATTGGAGTCTGCACAAAAACATCCATTTAATTGATTATGAGAGCTAGTG 2315
Qy      1779 CAGGGGGGCGACTTCTCTGACCTTCTCTCGCAGAGAGGGGCCCGCTCGGGGTGAAGACT 1838
Db      2316 CAAGGTGGAGACTTTTCAGACGTTTCTACAGAAATGAAGGTCCAGGCTGAAGGTGAAGAAG 2375
Qy      1839 CTGCTGCAGATGTGGGGGATGCAAGTGTGTCATGGAATGACTGAGAGCAAGTGTGC 1898
Db      2376 CTTATTAGGGTGTCAAAAAATGCTGCGGCTGGAATGGAATTTGGAGATGAACATTGT 2435
Qy      1899 ATCCACGGGACCTGCGCTGCTCGGAATGCTGCTGTGACAGAGAAAGATGCTTGAAGATC 1958
Db      2436 ATTCACAGGACCTTGACAGCCCGTAACTGTGTGTCATGAGAAAATGCTTAAAGATA 2495
Qy      1959 AGTGACTTTGGAGATGCTCCGAGAGAGAAAGCGATGGGGTCTATGCAAGCTCAGGGGGCTC 2018
Db      2496 AGTGACTTTGGCATGTCCAGAGAAAGAAAGATGGAGTATTCATCCACTGGGGGGATG 2555
Qy      2019 AGACAAGTCCCGTGAAATGACCCGCACTGAAGCCCTTAACTAAGCCGCTACTCTCC 2078
Db      2556 AAGCAGATTCTTATTAAATGACGGCGCCAGAAAGCCCTCAATTACGGTGGTACAGTTCA 2615
Qy      2079 GAAAGGACGTGTGAGGCTTTGGCATCTGCTGTGGAGAACCTTCAGCCTGGGGGCTCC 2138
Db      2616 GAGAGTGAATGTTGGAGTTTGGATCTTACTGTGGAAAGCCTTCAGCTTGGGTTTGT 2675
Qy      2139 CCTATCCCACTCAGCAATCAGACAGACACGGAGTTTGTGAGAAAGGGGGCCGCTTG 2198
Db      2676 CTTATGACAGCTATGACCAACAGCAAGGAGCCATAGAACAGGGAGTTGGTCTG 2735
Qy      2199 CCCTGCCAGAGCTGTCTCTGATGCTGTTCAAGGCTCATGAGCAGTGTGGGCTAT 2258
Db      2736 CTTGTCCCGATPACTGCCCTGATGAAGTCTATAGCTTGTGATGCTGAGATGCTGGGAATAT 2795
Qy      2259 GAGCCTGGGCGGCGCCAGCTTCAGACACATCTACAGAGAGCTGACAGACATCCGAAG 2318
Db      2796 GATCCAAAGAGGGGCTTAATTTCAGCATTTGCAATCAGGTGCTGTCAACATCCGCAAG 2855
Qy      2319 CGGCATCGGTGAGGCT 2334
Db      2856 AGATACCGATAAGTCT 2871
```

Search completed: March 19, 2005, 10:13:33
Job time : 1156.3 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 15:59:37 ; Search time 144 Seconds

(without alignments)
1726.159 Million cell updates/sec

Title: US-10-660-763-2

Perfect score: 3879

Sequence: 1 MGFSSSLCSPOGHVLOQM.....QRPSTIYQLSIRKXHR 752

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1401741 seqs, 330541175 residues

Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3879	100.0	752	13	US-10-003-295-2
2	3879	100.0	752	15	US-10-660-763-2
3	3834	98.8	822	13	US-10-003-295-4
4	3834	98.8	822	15	US-10-660-763-4
5	1853	47.8	822	9	US-09-886-319A-64
6	1853	47.8	822	14	US-10-376-556-64
7	1829.5	47.2	823	9	US-09-886-319A-63
8	1829.5	47.2	823	14	US-10-376-556-63
9	1528	39.4	292	16	US-10-664-421-69
10	1444	37.2	275	16	US-10-763-418-38
11	911	23.5	251	15	US-10-464-805-30
12	899	23.2	251	8	US-08-987-689A-30
13	899	23.2	251	14	US-10-292-524-30

14	669.5	17.3	257	10	US-09-823-187-46	Sequence 46, Appl
15	668.5	17.2	257	11	US-09-863-776-42	Sequence 42, Appl
16	661	17.0	258	11	US-09-964-956-75	Sequence 75, Appl
17	661	17.0	258	15	US-10-042-865-100	Sequence 100, App
18	661	17.0	258	15	US-10-072-012-801	Sequence 801, App
19	661	17.0	258	15	US-10-072-012-858	Sequence 858, App
20	657	16.9	256	10	US-09-976-782-40	Sequence 29, Appl
21	657	16.9	257	10	US-09-976-782-39	Sequence 40, Appl
22	657	16.9	257	15	US-10-087-684-100	Sequence 100, App
23	657	16.9	257	15	US-10-218-779-100	Sequence 100, App
24	656	16.9	254	15	US-10-052-648A-43	Sequence 43, Appl
25	655	16.9	2029	16	US-10-746-545-33	Sequence 33, Appl
26	642	16.6	250	15	US-10-074-978A-153	Sequence 153, App
27	642	16.6	280	9	US-09-515-806-15	Sequence 15, Appl
28	642	16.6	280	15	US-10-423-543-17	Sequence 17, Appl
29	624.5	16.1	1130	14	US-10-204-041-4	Sequence 4, Appl
30	624.5	16.1	1148	15	US-10-170-385-181	Sequence 181, App
31	624.5	16.1	1148	16	US-10-408-765A-746	Sequence 746, App
32	623.5	16.1	278	14	US-10-172-088-9	Sequence 9, Appl
33	623	16.1	244	15	US-10-042-865-85	Sequence 85, Appl
34	623	16.1	1123	16	US-10-746-545-10	Sequence 10, Appl
35	623	16.1	1123	16	US-10-746-545-11	Sequence 11, Appl
36	623	16.1	1149	15	US-10-457-954-6	Sequence 6, Appl
37	622	16.0	1130	14	US-10-171-889-1	Sequence 1, Appl
38	622	16.0	1130	14	US-10-276-633-4	Sequence 4, Appl
39	622	16.0	1130	14	US-10-263-480-2	Sequence 2, Appl
40	620.5	16.0	1567	14	US-10-312-918-2	Sequence 2, Appl
41	620.5	16.0	1594	14	US-10-312-918-4	Sequence 4, Appl
42	618	15.9	343	16	US-10-664-421-57	Sequence 57, Appl
43	614.5	15.8	537	16	US-10-746-545-12	Sequence 12, Appl
44	614.5	15.8	537	16	US-10-746-545-13	Sequence 13, Appl
45	614.5	15.8	537	16	US-10-746-545-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-10-003-295-2

Sequence 2, Application US/10003295

Publication No. US20020168741A1

GENERAL INFORMATION:

APPLICANT: GAN, Weiniu et al.

TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

FILE REFERENCE: THERMOF

FILE REFERENCE: CL001183DIY

CURRENT APPLICATION NUMBER: US/10/003,295

CURRENT FILING DATE: 2001-12-06

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 2

LENGTH: 752

TYPE: PRT

ORGANISM: Homo sapiens

US-10-003-295-2

Query Match	100.0%	Score 3879	DB 13	Length 752
Best Local Similarity	100.0%	Pred. No. 2.8e-223		
Matches 752	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MGFSSSLCSPOGHVLOQM	OEAEELRLBGRKMAQ	ORVKSDERYAGLLHMSLDSGGOS 60
DB	1	MGFSSSLCSPOGHVLOQM	OEAEELRLBGRKMAQ	ORVKSDERYAGLLHMSLDSGGOS 60
QY	61	RAISPSPISOSWARTSQTEGLSRLLRHOARDLNSGPISKLILIRERQOLRKYSEOM 120		
DB	61	RAISPSPISOSWARTSQTEGLSRLLRHOARDLNSGPISKLILIRERQOLRKYSEOM 120		
QY	121	QOLQOELTTHSODIEKLKSYRALARDSQAQKRYQASKODKDKAKDKVRSIMKLF 180		
DB	121	QOLQOELTTHSODIEKLKSYRALARDSQAQKRYQASKODKDKAKDKVRSIMKLF 180		

QY 181 AHHNRVYLGVRRAQLHGHQHOLLPGILRSLODLHEEMACILKEILOEYLEISLVODE 240
DB 181 AHHNRVYLGVRRAQLHGHQHOLLPGILRSLODLHEEMACILKEILOEYLEISLVODE 240
QY 241 VVAIHREMAAAARIQPEAEYOGFLROYGSAADVPDPCVTFDSLSLEEGPBLEGELQINE 300
DB 241 VVAIHREMAAAARIQPEAEYOGFLROYGSAADVPDPCVTFDSLSLEEGPBLEGELQINE 300
QY 301 LVESVQHTLTSVTBDELAVALATEMVFRRQEMVYQLOQELNBEENHTPRRVOVLGKROYL 360
DB 301 LVESVQHTLTSVTBDELAVALATEMVFRRQEMVYQLOQELNBEENHTPRRVOVLGKROYL 360
QY 361 QEALQGLQVALCSQAKLQAOQELLOTKLEHLGPBEPVYLLLODDHSTSSSQBEREGR 420
DB 361 QEALQGLQVALCSQAKLQAOQELLOTKLEHLGPBEPVYLLLODDHSTSSSQBEREGR 420
QY 421 TPTLEILKSHISGIFRPKPSNLYRLEGEFPISILLIDHLSTQOPLTKSGVVLHRAVP 480
DB 421 TPTLEILKSHISGIFRPKPSNLYRLEGEFPISILLIDHLSTQOPLTKSGVVLHRAVP 480
QY 481 KKKWVLNHDVLVGEQIGRGNFGEVFSGRLRADNTLVAVKSCRETLPPDLKAKFLQEAR1 540
DB 481 KKKWVLNHDVLVGEQIGRGNFGEVFSGRLRADNTLVAVKSCRETLPPDLKAKFLQEAR1 540
QY 541 LKQYSHPNIVRLIGVCTOKOPYIYIMELVQGGDFLFLTBEGARLRVKTLLQWVGDAAG 600
DB 541 LKQYSHPNIVRLIGVCTOKOPYIYIMELVQGGDFLFLTBEGARLRVKTLLQWVGDAAG 600
QY 601 MEYLSKCCIHDLAARNCLVTEKNVLIKISDFGMSREADGVYAASGGLROYVVKWTABE 660
DB 601 MEYLSKCCIHDLAARNCLVTEKNVLIKISDFGMSREADGVYAASGGLROYVVKWTABE 660
QY 661 ALNYGRYSESDVWSFGILLMETFSLGASPYNLSNOQTRFEVEKGRLLPCBELCPDAVF 720
DB 661 ALNYGRYSESDVWSFGILLMETFSLGASPYNLSNOQTRFEVEKGRLLPCBELCPDAVF 720
QY 721 RLMEQCWAYEPGQRPFSFTIYOELQSIKRRHR 752
DB 721 RLMEQCWAYEPGQRPFSFTIYOELQSIKRRHR 752

RESULT 2

US-10-660-763-2
; Sequence 2, Application US/10660763
; Publication No. US20040063130A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001183DIYI
; CURRENT APPLICATION NUMBER: US/10/660,763
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 752
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-660-763-2

Query Match 100.0%; Score 3879; DB 15; Length 752;
Best Local Similarity 100.0%; Pred. No. 2,8e-223;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSSSELCSPOGHGVLQOMQEAELRLLEGMRKMAORVXSDBREYAGLLHMSLODSGOS 60
DB 1 MGSSSELCSPOGHGVLQOMQEAELRLLEGMRKMAORVXSDBREYAGLLHMSLODSGOS 60
QY 61 RAISPSPIISQSAEITSGTEGSLRLRQHAEDLNSGPKSLLLRROQLKTTYSEOW 120
DB 61 RAISPSPIISQSAEITSGTEGSLRLRQHAEDLNSGPKSLLLRROQLKTTYSEOW 120

QY 121 QOLOQELTKTSHQDIEKLSQYRALARDSAQAKRYQOASXKXDRDKAKDYVSLMKLF 180
DB 121 QOLOQELTKTSHQDIEKLSQYRALARDSAQAKRYQOASXKXDRDKAKDYVSLMKLF 180
QY 181 AHHNRVYLGVRRAQLHGHQHOLLPGILRSLODLHEEMACILKEILOEYLEISLVODE 240
DB 181 AHHNRVYLGVRRAQLHGHQHOLLPGILRSLODLHEEMACILKEILOEYLEISLVODE 240
QY 241 VVAIHREMAAAARIQPEAEYOGFLROYGSAADVPDPCVTFDSLSLEEGPBLEGELQINE 300
DB 241 VVAIHREMAAAARIQPEAEYOGFLROYGSAADVPDPCVTFDSLSLEEGPBLEGELQINE 300
QY 301 LVESVQHTLTSVTBDELAVALATEMVFRRQEMVYQLOQELNBEENHTPRRVOVLGKROYL 360
DB 301 LVESVQHTLTSVTBDELAVALATEMVFRRQEMVYQLOQELNBEENHTPRRVOVLGKROYL 360
QY 361 QEALQGLQVALCSQAKLQAOQELLOTKLEHLGPBEPVYLLLODDHSTSSSQBEREGR 420
DB 361 QEALQGLQVALCSQAKLQAOQELLOTKLEHLGPBEPVYLLLODDHSTSSSQBEREGR 420
QY 421 TPTLEILKSHISGIFRPKPSNLYRLEGEFPISILLIDHLSTQOPLTKSGVVLHRAVP 480
DB 421 TPTLEILKSHISGIFRPKPSNLYRLEGEFPISILLIDHLSTQOPLTKSGVVLHRAVP 480
QY 481 KKKWVLNHDVLVGEQIGRGNFGEVFSGRLRADNTLVAVKSCRETLPPDLKAKFLQEAR1 540
DB 481 KKKWVLNHDVLVGEQIGRGNFGEVFSGRLRADNTLVAVKSCRETLPPDLKAKFLQEAR1 540
QY 541 LKQYSHPNIVRLIGVCTOKOPYIYIMELVQGGDFLFLTBEGARLRVKTLLQWVGDAAG 600
DB 541 LKQYSHPNIVRLIGVCTOKOPYIYIMELVQGGDFLFLTBEGARLRVKTLLQWVGDAAG 600
QY 601 MEYLSKCCIHDLAARNCLVTEKNVLIKISDFGMSREADGVYAASGGLROYVVKWTABE 660
DB 601 MEYLSKCCIHDLAARNCLVTEKNVLIKISDFGMSREADGVYAASGGLROYVVKWTABE 660
QY 661 ALNYGRYSESDVWSFGILLMETFSLGASPYNLSNOQTRFEVEKGRLLPCBELCPDAVF 720
DB 661 ALNYGRYSESDVWSFGILLMETFSLGASPYNLSNOQTRFEVEKGRLLPCBELCPDAVF 720
QY 721 RLMEQCWAYEPGQRPFSFTIYOELQSIKRRHR 752
DB 721 RLMEQCWAYEPGQRPFSFTIYOELQSIKRRHR 752

RESULT 3

US-10-003-295-4
; Sequence 4, Application US/10003295
; Publication No. US20020168741A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001183DIYI
; CURRENT APPLICATION NUMBER: US/10/003,295
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-003-295-4

Query Match 98.8%; Score 3834; DB 13; Length 822;
Best Local Similarity 91.5%; Pred. No. 1,5e-220;
Matches 752; Conservative 0; Mismatches 0; Indels 70; Gaps 1;

QY 1 MGSSSELCSPOGHGVLQOMQEAELRLLEGMRKMAORVXSDBREYAGLLHMSLODSGOS 60
DB 1 MGSSSELCSPOGHGVLQOMQEAELRLLEGMRKMAORVXSDBREYAGLLHMSLODSGOS 60

```

QY 61 RAISPDSPISQSWAEITTSQTEGLSRLLRQHAEDLNSGPKSLRLRERQOLKTYSEOW 120
DB 61 RAISPDSPISQSWAEITTSQTEGLSRLLRQHAEDLNSGPKSLRLRERQOLKTYSEOW 120
QY 121 QOLQOELTTHSGDIEKLKSOYRALARDSQAQKRYQASKODKDRKADKQVRSIMKLF 180
DB 121 QOLQOELTTHSGDIEKLKSOYRALARDSQAQKRYQASKODKDRKADKQVRSIMKLF 180
QY 181 AHNHRYVLGVRAAQLHQNHHQHLPLGLRSLQDLHEBNACILKEILOEYLEISLVODE 240
DB 181 AHNHRYVLGVRAAQLHQNHHQHLPLGLRSLQDLHEBNACILKEILOEYLEISLVODE 240
QY 241 VVAIHREMAAARARQPEAEYQGLRQYGSAPDVPCTPDSLEBEGPBLEGELOLNE 300
DB 241 VVAIHREMAAARARQPEAEYQGLRQYGSAPDVPCTPDSLEBEGPBLEGELOLNE 300
QY 301 LTVESVQHTLTSVTDELAATEMVFRRQEMVTOLOQELNBEENTHPRERVOILGKROYL 360
DB 301 LTVESVQHTLTSVTDELAATEMVFRRQEMVTOLOQELNBEENTHPRERVOILGKROYL 360
QY 361 QEALQGLQVALCSQAQLQAOQELQTKLEHLGPEPPVULLQDDHSTSSSEQEREGGR 420
DB 361 QEALQGLQVALCSQAQLQAOQELQTKLEHLGPEPPVULLQDDHSTSSSEQEREGGR 420
QY 421 TPTLEILKSHISGIFRPKFS----- 440
DB 421 TPTLEILKSHISGIFRPKFS----- 440
QY 441 -----NLRYLEGEPPSIPLLIDHLLSTQOPLTKK 470
DB 441 LVBESQKQOEYVLSVMDGLPRHFIITQSLDNLRYLEGEPPSIPLLIDHLLSTQOPLTKK 470
QY 471 SGVYLHRAVPKQKMWLNHEDVLGEOIGRNGEVSFGRLRADNTLVAAKSCRETLPPDL 530
DB 471 SGVYLHRAVPKQKMWLNHEDVLGEOIGRNGEVSFGRLRADNTLVAAKSCRETLPPDL 530
QY 531 KAKFLOEABILKQYSHPNIVRLIGVCTQKOPYIYVLELVQGGDFLFLRTGEGARLKVTL 590
DB 531 KAKFLOEABILKQYSHPNIVRLIGVCTQKOPYIYVLELVQGGDFLFLRTGEGARLKVTL 590
QY 591 LQWVGDAAGMEYLESKCCIHRLDARNCIVTEKNVLTISDFGMSREBAGVYAAAGGLR 650
DB 591 LQWVGDAAGMEYLESKCCIHRLDARNCIVTEKNVLTISDFGMSREBAGVYAAAGGLR 650
QY 651 QVPVKMTABEALNYGYSSESVDVMSFGILLMETFSLGASPYPLSNQOTREFVEKGRRLP 710
DB 651 QVPVKMTABEALNYGYSSESVDVMSFGILLMETFSLGASPYPLSNQOTREFVEKGRRLP 710
QY 711 CPELCPDAVPRIMEQCMAYEPGQRPSPSTTYOELQSIIRKRRH 822
DB 711 CPELCPDAVPRIMEQCMAYEPGQRPSPSTTYOELQSIIRKRRH 822

RESULT 4
US-10-660-763-4
; Sequence 4, Application US/10660763
; Publication No. US20040063130A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001183DIV1
; CURRENT APPLICATION NUMBER: US/10/660,763
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-660-763-4

```

```

Query Match      98.8%; Score 3834; DB 15; Length 822;
Best Local Similarity 91.5%; Pred. No. 1,5e-220;
Matches 752; Conservative 0; Mismatches 0; Indels 70; Gaps 1;

QY 1 MGFSSSELCPQHGVLQOQOEAELRLLEGARKMAQRYVSDREYAGLLHMSIQDSGQGS 60
DB 1 MGFSSSELCPQHGVLQOQOEAELRLLEGARKMAQRYVSDREYAGLLHMSIQDSGQGS 60
QY 61 RAISPDSPISQSWAEITTSQTEGLSRLLRQHAEDLNSGPKSLRLRERQOLKTYSEOW 120
DB 61 RAISPDSPISQSWAEITTSQTEGLSRLLRQHAEDLNSGPKSLRLRERQOLKTYSEOW 120
QY 121 QOLQOELTTHSGDIEKLKSOYRALARDSQAQKRYQASKODKDRKADKQVRSIMKLF 180
DB 121 QOLQOELTTHSGDIEKLKSOYRALARDSQAQKRYQASKODKDRKADKQVRSIMKLF 180
QY 181 AHNHRYVLGVRAAQLHQNHHQHLPLGLRSLQDLHEBNACILKEILOEYLEISLVODE 240
DB 181 AHNHRYVLGVRAAQLHQNHHQHLPLGLRSLQDLHEBNACILKEILOEYLEISLVODE 240
QY 241 VVAIHREMAAARARQPEAEYQGLRQYGSAPDVPCTPDSLEBEGPBLEGELOLNE 300
DB 241 VVAIHREMAAARARQPEAEYQGLRQYGSAPDVPCTPDSLEBEGPBLEGELOLNE 300
QY 301 LTVESVQHTLTSVTDELAATEMVFRRQEMVTOLOQELNBEENTHPRERVOILGKROYL 360
DB 301 LTVESVQHTLTSVTDELAATEMVFRRQEMVTOLOQELNBEENTHPRERVOILGKROYL 360
QY 361 QEALQGLQVALCSQAQLQAOQELQTKLEHLGPEPPVULLQDDHSTSSSEQEREGGR 420
DB 361 QEALQGLQVALCSQAQLQAOQELQTKLEHLGPEPPVULLQDDHSTSSSEQEREGGR 420
QY 421 TPTLEILKSHISGIFRPKFS----- 440
DB 421 TPTLEILKSHISGIFRPKFS----- 440
QY 441 -----NLRYLEGEPPSIPLLIDHLLSTQOPLTKK 470
DB 441 LVBESQKQOEYVLSVMDGLPRHFIITQSLDNLRYLEGEPPSIPLLIDHLLSTQOPLTKK 470
QY 471 SGVYLHRAVPKQKMWLNHEDVLGEOIGRNGEVSFGRLRADNTLVAAKSCRETLPPDL 530
DB 471 SGVYLHRAVPKQKMWLNHEDVLGEOIGRNGEVSFGRLRADNTLVAAKSCRETLPPDL 530
QY 531 KAKFLOEABILKQYSHPNIVRLIGVCTQKOPYIYVLELVQGGDFLFLRTGEGARLKVTL 590
DB 531 KAKFLOEABILKQYSHPNIVRLIGVCTQKOPYIYVLELVQGGDFLFLRTGEGARLKVTL 590
QY 591 LQWVGDAAGMEYLESKCCIHRLDARNCIVTEKNVLTISDFGMSREBAGVYAAAGGLR 650
DB 591 LQWVGDAAGMEYLESKCCIHRLDARNCIVTEKNVLTISDFGMSREBAGVYAAAGGLR 650
QY 651 QVPVKMTABEALNYGYSSESVDVMSFGILLMETFSLGASPYPLSNQOTREFVEKGRRLP 710
DB 651 QVPVKMTABEALNYGYSSESVDVMSFGILLMETFSLGASPYPLSNQOTREFVEKGRRLP 710
QY 711 CPELCPDAVPRIMEQCMAYEPGQRPSPSTTYOELQSIIRKRRH 822
DB 711 CPELCPDAVPRIMEQCMAYEPGQRPSPSTTYOELQSIIRKRRH 822

RESULT 5
US-09-886-319A-64
; Sequence 64, Application US/09886319A
; Publication No. US20020086019A1
; GENERAL INFORMATION:
; APPLICANT: Wolf, Eckard
; APPLICANT: Weimer, Sabine
; APPLICANT: Halle, Jörn-Peter
; APPLICANT: Regembogen, Johannes
; APPLICANT: Goppelt, Andreas
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound

```

```

/ TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
/ FILE OF INVENTION: Active Substances
/ FILE REFERENCE: 50125/014002
/ CURRENT APPLICATION NUMBER: US/09/886,319A
/ CURRENT FILING DATE: 2001-06-20
/ PRIOR APPLICATION NUMBER: US 60/222,081
/ PRIOR FILING DATE: 2000-08-01
/ PRIOR APPLICATION NUMBER: DE 10030149.5
/ PRIOR FILING DATE: 2000-06-20
/ NUMBER OF SEQ ID NOS: 84
/ SOFTWARE: FaalSeq for Windows Version 4.0
/ SEQ ID NO 64
/ LENGTH: 822
/ TYPE: PR
/ ORGANISM: Homo sapiens
US-09-886-319A-64

Query Match      47.8%; Score 1853; DB 9; Length 822;
Best Local Similarity 45.5%; Pred. No. 2,6e-102;
Matches 376; Conservative 150; Mismatches 217; Indels 84; Gaps 9;

QY 1 MGFSSSLCSPQCHGVLCQOQAEALRLLEGMRKMAQRYVSDRYAGLHHMSLQDSGGGS 60
DB 1 MGFSSDL--KNSHEAVLKQDWELRLLETVKFMALRISDKEXASTLQNLQNVDKEST 58
QY 61 RAISPDSPISQSWARTSOTEGLSRLRQHAEDLNSGPISKLILRERQOLKTYSEQM 120
DB 59 VQMNIVSNVSKSWMLIQOTEQLSRIMKTHAEDLNSGPLRLTMMKDKQVKKSTIGVH 118
QY 121 QOLQOELTGTSHODIEKLKSQYRALARDSQAQRYQEA-SKODRKAKDKYVRSIMTL 179
DB 119 QOLIEAMIKVTTELKELKCSYROLIKEMNSAKETKALAKETEKAKERYDKATMKL 178
QY 180 FAHNRVYLVGVAQOLHONHOLLPGLLRSLODHEEMACILKEILOEYELISLVOD 239
DB 179 HMLHQYVLTALKAQOLHONQYDITPLRLDLSQKQOEBIMIALKGIPEYEQITSLYTE 238
QY 240 EVVAIIRREMAAARLQPEAEYQGFLOYGASAPDVPCPTFDESLEEGEPLEPGLQIN 299
DB 239 EIVNVHKEIQMSVEQIDPSTEYNNFTDVHRTTAKEQOELEFDTLSLEENENLOANEIMN 298
QY 300 ELTVESVQHTLTSVDELAVATEMVRROEMVTOLOEELRNEENTHPR-ERVOLLGKRQ 358
DB 299 NITAEISLQWMLKTAEELMQOQMLNKEAVLELEKRIEESSETECKSDIVLLSQKQ 358
QY 359 VLOEALQGLQVALCSQAQKLAQOELLQTLKLEHNGPPEPPVLLLODRHSTSSSEBERG 418
DB 359 ALEELKQSVQQLRCTEAKFSAQKELLEQKVQENDGKEPPVNVYEDDASVTSMERKE-- 416
QY 419 GRTPLTEILKSHISGIFR-PK----- 438
DB 417 -RLSKFESTIRHSIAGIIRSPKAVGSSALSMDMSISEKPLAEDQWYHGAIPRIEAOELLK 475
QY 439 -----FSNLVRLBEGGEPSPISPLIDHLLSTQ 464
DB 476 KQGDPLVRBSHGKRGXYLVSVSDGQRHPIIQVYNNMVRFBEGTGSNIPQLIDHHTTK 535
QY 465 QPITTKSGVYLHAAYKRD-KVTLNHEDELVLGEQIGNGNGEVSGLRADNTLVAVKSCR 523
DB 536 QVITTKSGVYLLNPIPRDKKMLISHEDVILGELLAGNGFGEVYKGLTK-DKTSVAAYTKK 594
QY 524 ETLRPLTKAKFLQOEAELIKQYSHNIVRLIGVCTOKOPITYIWEILOVQSGPLFTLREG 583
DB 595 EDLPQELKIFLQDEAKILKQYDHPNIVKLIGVCTQOPVYIIMELVSGGDFLFLRRKQD 654
QY 584 RLKRVKTLQWVGDAAGMEYLESKSCIRHDLAARNCLVTEKNVLTISDFGMSREEDAVY 643
DB 655 ELKTLQVLRKSLDAAAGMLYLESKNCIHRDLAARNCLVGENNVLTISDFGMSQDEGQYV 714
QY 644 AASGGIRQVYVKTABALNLYGRYSSESDYWSFGIILMTFSGASPYRNLNSQOTREYV 703
DB 715 SSS-GLKOPIPIKTABALNLYGRYSSESDYWSFGIILMTFSGVCPYEGMTNQOAREOV 773

```

```

QY 704 EKGRLPCPELCPDAVFRIMEQWMAYPEQRPSTIYLOELQIRKR 750
DB 774 ERGYMSAPQHCPEDISKIMMKCMWYKPPENRKFSFLOKELVITIKK 820

RESULT 6
US-10-376-564-64
/ Sequence 64, Application US/10376564
/ Publication No. US20030180302A1
/ GENERAL INFORMATION:
/ APPLICANT: Wolf, Eckhard
/ APPLICANT: Werner, Sabine
/ APPLICANT: Haile, Jörn-Peter
/ APPLICANT: Regenhagen, Johannes
/ APPLICANT: Goppelt, Andreas
/ TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
/ TITLE OF INVENTION: the diagnosis or treatment of skin disorders and wound
/ TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
/ FILE REFERENCE: 50125/014003
/ CURRENT APPLICATION NUMBER: US/10/376,564
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 09/886,319
/ PRIOR FILING DATE: 2001-06-20
/ PRIOR APPLICATION NUMBER: US 60/222,081
/ PRIOR FILING DATE: 2000-08-01
/ PRIOR APPLICATION NUMBER: DE 100 30 149.5-41
/ PRIOR FILING DATE: 2000-06-20
/ NUMBER OF SEQ ID NOS: 86
/ SOFTWARE: FaalSeq for Windows Version 4.0
/ SEQ ID NO 64
/ LENGTH: 822
/ TYPE: PR
/ ORGANISM: Homo sapiens
US-10-376-564-64

Query Match      47.8%; Score 1853; DB 14; Length 822;
Best Local Similarity 45.5%; Pred. No. 2,6e-102;
Matches 376; Conservative 150; Mismatches 217; Indels 84; Gaps 9;

QY 1 MGFSSSLCSPQCHGVLCQOQAEALRLLEGMRKMAQRYVSDRYAGLHHMSLQDSGGGS 60
DB 1 MGFSSDL--KNSHEAVLKQDWELRLLETVKFMALRISDKEXASTLQNLQNVDKEST 58
QY 61 RAISPDSPISQSWARTSOTEGLSRLRQHAEDLNSGPISKLILRERQOLKTYSEQM 120
DB 59 VQMNIVSNVSKSWMLIQOTEQLSRIMKTHAEDLNSGPLRLTMMKDKQVKKSTIGVH 118
QY 121 QOLQOELTGTSHODIEKLKSQYRALARDSQAQRYQEA-SKODRKAKDKYVRSIMTL 179
DB 119 QOLIEAMIKVTTELKELKCSYROLIKEMNSAKETKALAKETEKAKERYDKATMKL 178
QY 180 FAHNRVYLVGVAQOLHONHOLLPGLLRSLODHEEMACILKEILOEYELISLVOD 239
DB 179 HMLHQYVLTALKAQOLHONQYDITPLRLDLSQKQOEBIMIALKGIPEYEQITSLYTE 238
QY 240 EVVAIIRREMAAARLQPEAEYQGFLOYGASAPDVPCPTFDESLEEGEPLEPGLQIN 299
DB 239 EIVNVHKEIQMSVEQIDPSTEYNNFTDVHRTTAKEQOELEFDTLSLEENENLOANEIMN 298
QY 300 ELTVESVQHTLTSVDELAVATEMVRROEMVTOLOEELRNEENTHPR-ERVOLLGKRQ 358
DB 299 NITAEISLQWMLKTAEELMQOQMLNKEAVLELEKRIEESSETECKSDIVLLSQKQ 358
QY 359 VLOEALQGLQVALCSQAQKLAQOELLQTLKLEHNGPPEPPVLLLODRHSTSSSEBERG 418
DB 359 ALEELKQSVQQLRCTEAKFSAQKELLEQKVQENDGKEPPVNVYEDDASVTSMERKE-- 416
QY 419 GRTPLTEILKSHISGIFR-PK----- 438
DB 417 -RLSKFESTIRHSIAGIIRSPKAVGSSALSMDMSISEKPLAEDQWYHGAIPRIEAOELLK 475
QY 439 -----FSNLVRLBEGGEPSPISPLIDHLLSTQ 464

```



```

Db      476 KQGDPLVRESHKRGBYLVSYSDGQRHFIIOYDMMTRFEGTGSNIPOLIDHHYTK 535
Qy      465 QPLTKKSGVLIHRAVPKD-KVNLNHEDLVLGEQIGRNGFGEVFSGLRADNTLVAVKSCR 523
Db      536 QVTTKSGVLIHRAVPKD-KVNLNHEDLVLGEQIGRNGFGEVFSGLRADNTLVAVKSCR 594
Qy      524 ETLPRDLKAKFLQEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLFLRTG 583
Db      595 EDLPQELKFLQEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLFLRTG 654
Qy      584 RLKVKTLQWVGDAAGMEYLESKCCIRHDLAARNCILYETKNVLKISDFGMSREBAGVY 643
Db      655 ELKQLVLFPSDLAAGMILYLESKCCIRHDLAARNCILYETKNVLKISDFGMSREBAGVY 714
Qy      644 AASGGLQVVPVKTABEALNRYGSSSDVMSFGILLMETFSLGASPYPNLSNOQTRFV 703
Db      715 SSS-GLKQIPKWTABEALNRYGSSSDVMSFGILLMETFSLGASPYPNLSNOQTRFV 773
Qy      704 EKGGRLPCELCDAVFRMLMEQCMAYEPGQRPFSFTIYQELQSIK 750
Db      774 EKGGRMSAPQHCPEBIDISKIMKCMYKPEKRPFSFELQELTIK 820

```

```

RESULT 7
US-09-886-319A-63
/ Sequence 63, Application US/09886319A
/ Publication No. US20020086019A1
/ GENERAL INFORMATION:
/ APPLICANT: Wolf, Eckard
/ APPLICANT: Werner, Sabine
/ APPLICANT: Halle, Jörn-Peter
/ APPLICANT: Regenbogen, Johannes
/ APPLICANT: Goppelt, Andreas
/ TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
/ TITLE OF INVENTION: the diagnosis or Treatment of Skin Disorders and Wound
/ TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
/ TITLE OF INVENTION: Active Substances
/ FILE REFERENCE: 50125/014002
/ CURRENT APPLICATION NUMBER: US/09/886,319A
/ CURRENT FILING DATE: 2001-06-20
/ PRIOR APPLICATION NUMBER: US 60/222,081
/ PRIOR FILING DATE: 2000-08-01
/ PRIOR APPLICATION NUMBER: DE 10030149.5
/ PRIOR FILING DATE: 2000-06-20
/ NUMBER OF SEQ ID NOS: 84
/ SOFTWARE: FaastSeq for Windows Version 4.0
/ SEQ ID NO 63
/ LENGTH: 823
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-09-886-319A-63

```

```

Query Match      47.2%; Score 1829.5; DB 9; Length 823;
Best Local Similarity 45.6%; Pred. No. 6.5e-101;
Matches 377; Conservative 146; Mismatches 219; Indels 85; Gaps 10;

```

```

Qy      1 MGFSSSLCSPOHGVLCQOAEALRLLEGKMKMAQRYVSDREYAGILHMSLQDSGGS 60
Db      1 MEGGSLKNSQ-BAYL-KIQDWELRLLETKKFMALRISDKEXAYTLQNLQNVKEST 58
Qy      61 RAISPSPISQSWAETSTQTEGSLRLRQHAEDLNSGRLSKSLIRERQRLKTYSEGM 120
Db      59 VQNVYVSNYSKSVLMIQOTEBLSRLMKTAEDLNSGRLHRLTWMIKDQOVKKSIVGJH 118
Qy      121 OOLQOELTTHSDIKLSQVYALARDASQAKRYQEA-SKQKDXDKAKDYVRSMLK 179
Db      119 QOLBAEMIKYTKTELEKSSYRQLTKENMSAKETKEALAKETKAKERYDKATMK 178
Qy      180 FAHNNRYVLGVRAAQALHQNHLPLGLRLSLQDHEEMACILKEILOEYLISSLYVD 239
Db      179 HMLHNGYVALAKGAQHSGYVDTTLPLLLDSVQKQOENIKALKGIPDDYSQITSLVTE 238

```

```

Qy      240 EVVAIHRMAAAARIQEAERYOGFLROYGSAVDVPCVTFPESSLEEGEPLEPGLQLN 299
Db      239 EIVNHKEIQMSVEQIDPSTETNNFIDVARTTAKEQIEIFDLSLEENLQANETMMN 298
Qy      300 ELTVESQHTLTSVTDLAVATEMFRROEMVTOLOEELRNEENHPR-ERVOLIGKRO 358
Db      299 NLTADSLQVMTLTIABELVQTOQMLHKAVALLEKRIEESPTECKKSDVILLGQNG 358
Qy      359 VLQELQIGVALCSQAKLQAOQELIQTKLBNLGRPEPPVLLDODRHSSTSSQOERBG 418
Db      359 ALBELKQSVQOQRCSAKCAQKALLEQYQENDKKEPPVNVYBEDARSVTSMEKE-- 416
Qy      419 GRTPTLEIKSHISGFR-PK----- 438
Db      417 -RLSKFESTIRHSIAGITKSPKSVLSSITQCVISYGERPLAEHDMTHCAIPRIEAOELL 475
Qy      439 -----FSNLYRLEGEGFPSPILPLIDHLST 463
Db      476 KQGDPLVRESHKRGBYLVSYSDGQRHFIIOYDMMTRFEGTGSNIPOLIDHHYTK 535
Qy      464 QPLTKKSGVLIHRAVPKD-KVNLNHEDLVLGEQIGRNGFGEVFSGLRADNTLVAVKSC 522
Db      536 QVTTKSGVLIHRAVPKD-KVNLNHEDLVLGEQIGRNGFGEVFSGLRADNTLVAVKSC 594
Qy      523 ETLPRDLKAKFLQEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLFLRTG 582
Db      595 EDLPQELKFLQEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLFLRTG 654
Qy      584 RLKVKTLQWVGDAAGMEYLESKCCIRHDLAARNCILYETKNVLKISDFGMSREBAGV 642
Db      655 ELKQLVLFPSDLAAGMILYLESKCCIRHDLAARNCILYETKNVLKISDFGMSREBAGV 714
Qy      644 AASGGLQVVPVKTABEALNRYGSSSDVMSFGILLMETFSLGASPYPNLSNOQTRFV 702
Db      715 SSS-GLKQIPKWTABEALNRYGSSSDVMSFGILLMETFSLGASPYPNLSNOQTRFV 773
Qy      704 EKGGRLPCELCDAVFRMLMEQCMAYEPGQRPFSFTIYQELQSIK 749
Db      774 EKGGRMSAPQHCPEBIDISKIMKCMYKPEKRPFSFELQELTIK 820

```

```

RESULT 8
US-10-376-564-63
/ Sequence 63, Application US/10376564
/ Publication No. US20030180302A1
/ GENERAL INFORMATION:
/ APPLICANT: Wolf, Eckhard
/ APPLICANT: Werner, Sabine
/ APPLICANT: Halle, Jörn-Peter
/ APPLICANT: Regenbogen, Johannes
/ APPLICANT: Goppelt, Andreas
/ TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
/ TITLE OF INVENTION: the diagnosis or Treatment of Skin Disorders and Wound
/ TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
/ TITLE OF INVENTION: Active Substances
/ FILE REFERENCE: 50125/014003
/ CURRENT APPLICATION NUMBER: US/10/376,564
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 09/886,319
/ PRIOR FILING DATE: 2001-06-20
/ PRIOR APPLICATION NUMBER: US 60/222,081
/ PRIOR FILING DATE: 2000-08-01
/ PRIOR APPLICATION NUMBER: DE 100 30 149.5-41
/ PRIOR FILING DATE: 2000-06-20
/ NUMBER OF SEQ ID NOS: 86
/ SOFTWARE: FaastSeq for Windows Version 4.0
/ SEQ ID NO 63
/ LENGTH: 823
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-10-376-564-63
Query Match      47.2%; Score 1829.5; DB 14; Length 823;

```

Best Local Similarity 45.6%; Pred. No. 6.5e-101;
Matches 377; Conservative 146; Mismatches 219; Indels 85; Gaps 10;

```

Qy 1 MGFSSSLCPQGHGVCVQOQZAEALLEGKMKMAQKQVSDREYAGLLHHMSLQDSGGGS 60
Db 1 MGFSGDLKNSQ-BAYL-KLQDWELRLLETVKKFMALRISDKEXAYATLQNLNQVDKEST 58
Qy 61 RAISPDSPISQSWAEITSGEGSLRLRQHAEDLNSGPKSLILRERQOLRKYSEGM 120
Db 59 VQVNVYNSVSKSLMLTQOTBOLSRIMKTHAEDLNSGPKSLILRERQOLRKYSEGM 118
Qy 121 QQLQOELTHTSHQDIEKLKSOYRALARDAQAQKRYQEA-SKDQKDAKDKYVSLMTL 179
Db 119 QQIEAMKIVTTEKLEKSSYRQLIKEMNSAKETKALAKGETEKAKERDKATMTL 178
Qy 180 FAHNRVYLVGAPAAQHLQHHLHQLLPGLIRSLQDHEMACLKILQEYLEISLVOD 239
Db 179 HMLHNOYVALKMAQHLQSOYDYDTLPRLILDSVQKQEMIKALKGIFFDYSQISLVT 238
Qy 240 EYVAIHRMAAARLOPEAEYQFLROYGSAADVPDPCVTPDESLEEGEPLEPELQLN 299
Db 239 EYVNHKEIOMSVQEDIDPSTBYNFIQVHRTTAAKQIEIFDTSLSEENLQANEIMN 298
Qy 300 ELTVESVQHTLTSVTDLAVALTEVFRQEMVTOLOQELRNEBENTHPR-ERYOLLGKQ 358
Db 299 NLTAQSLQVMTLKTALBELTQTOQMLHKEAVALLEKRIEESPETCEKSDIVLLGQK 358
Qy 359 VLOEALQGLQVALCSQAKIQAQOELLQTLKLEHGPBEPFVLLLODRSTSSSEBERG 418
Db 359 AAEELQSVQOALRCSAKAQAQALLERQVQENDKEPPVNVYEDASVYSMERKE-- 416
Qy 419 GRTPLTEIKSHISGIFR-PK----- 438
Db 417 -RLSKESIRHSIAGIKSPKSVLGSSTQVCVIVSGERPLAHDVYHGAIRIENQELL 475
Qy 439 -----FSNLYRLEGSGFPSPILILIDHLST 463
Db 476 KQGGDFLVRESHQKPEEYVLVSVDGRRHFIQFVDNLVRFEGTGFNSNIPQLIDHFMNT 535
Qy 464 QOPLTKKSGVVLHRAVPKQ-KVTLNHEDLVGEQIRGNFGEVFSRLADNTLVAVKSC 522
Db 536 KQVITTKSGSVLNLPIPKQKWLNLNEDVSLGELLKGNFGEVYKQTLK-DKTPVALIKTC 594
Qy 523 RETLPDLKAKFLQEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLTFELRT 582
Db 595 KEDLPBELIKFLQEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLTFELRT 654
Qy 583 ARLRVKTLLOMVGDAAGMEYLESKCCIHRLAARNCLVTEKNVLIKISDFGMSREBADV 642
Db 655 DELKTLQVRFSLDVAAGMLYLESKCCIHRLAARNCLVTEKNVLIKISDFGMSRQEDGV 714
Qy 643 YAAAGSLRQVPVMTPEALNRYRYSSESQVMSFGILLMETFSLGASPYNLSNOQTRF 702
Db 715 YSSS-GLKQIPIMTAPALNRYRYSSESQVMSFGILLMETFSLGASPYNLSNOQTRF 773
Qy 703 VEKGRGLPCPELCPDAVFRLEMCMAVEPQGRSFSFTIYQELQIRKR 749
Db 774 VERGYMSAPQNCPEVFTIMKCMYKPBENRKFPLDKELTVIYK 820

```

RESULT 9
US-10-664-421-69

```

; Sequence 69, Application US/10664421
; Publication No. US20040142864A1
; GENERAL INFORMATION:
; APPLICANT: BREMER, RYAN
; APPLICANT: IBRAHIM, PRABHA
; APPLICANT: KUMAR, ABHINAV
; APPLICANT: MANDIYAN, VALSAN
; APPLICANT: MILBURN, MICHAEL V.
; TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE
; FILE REFERENCE: 039363/0703
; CURRENT APPLICATION NUMBER: US/10/664,421

```

```

; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/412,341
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/411,398
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 69
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-664-421-69

```

Query Match 39.4%; Score 1528; DB 16; Length 292;
Best Local Similarity 99.7%; Pred. No. 2e-83;
Matches 291; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 461 LSTQOPLTKKSGVVLHRAVPKQKVLNHEDLVLSQIRGNFGEVFSGLRADNTLVAVK 520
Db 1 LSTQOPLTKKSGVVLHRAVPKQKVLNHEDLVLSQIRGNFGEVFSGLRADNTLVAVK 60
Qy 521 SRETLPLDKAKFLQEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLTFELRT 580
Db 61 SRETLPLDKAKFLQEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLTFELRT 120
Qy 581 EGARLRVKTLLQMVGDAAAGMEYLESKCCIHRLAARNCLVTEKNVLIKISDFGMSREB 640
Db 121 EGARLRVKTLLQMVGDAAAGMEYLESKCCIHRLAARNCLVTEKNVLIKISDFGMSREB 180
Qy 641 GYAAAGSLRQVPVMTPEALNRYRYSSESQVMSFGILLMETFSLGASPYNLSNOQTR 700
Db 181 GYAAAGSLRQVPVMTPEALNRYRYSSESQVMSFGILLMETFSLGASPYNLSNOQTR 240
Qy 701 EYVNHKEIOMSVQEDIDPSTBYNFIQVHRTTAAKQIEIFDTSLSEENLQANEIMN 752
Db 241 EYVNHKEIOMSVQEDIDPSTBYNFIQVHRTTAAKQIEIFDTSLSEENLQANEIMN 292

```

RESULT 10

```

US-10-763-418-38
; Sequence 38, Application US/10763418
; Publication No. US20040185547A1
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/10/763,418
; CURRENT FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: US/09/664,526
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/188,809
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-763-418-38

```

Query Match 37.2%; Score 1444; DB 16; Length 275;
Best Local Similarity 99.6%; Pred. No. 1.9e-78;
Matches 274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 478 AVPKQKVLNHEDLVLSQIRGNFGEVFSGLRADNTLVAVKSGRETLPLDKAKFLQ 537
Db 1 AVPKQKVLNHEDLVLSQIRGNFGEVFSGLRADNTLVAVKSGRETLPLDKAKFLQ 60

```

QY 538 ARILKOYSHPIVRLIGVCTOKOPITYIYVMEIVOGGDFLTFLRTGRLRVKTLQWVGA 597
| | | | |
DB 61 AAILKOYSHPIVRLIGVCTOKOPITYIYVMEIVOGGDFLTFLRTGRLRVKTLQWVGA 120
| | | | |
QY 598 AAGMEYLESKCIIHRDLAARNCVTEKYNLKIISDFGMSREADVAAAGGRLROVPVKT 657
| | | | |
DB 121 AAGMEYLESKCIIHRDLAARNCVTEKYNLKIISDFGMSREADVAAAGGRLROVPVKT 180
| | | | |
QY 658 APEALNAYGYSSESVDWVSGILMETFSIGASPYPLNSNOQTRFEVEKGRLLPCPELCPD 717
| | | | |
DB 181 APEALNAYGYSSESVDWVSGILMETFSIGASPYPLNSNOQTRFEVEKGRLLPCPELCPD 240
| | | | |
QY 718 AVFRLMEOCMAVEPGORPSFSTIYOELOSIRKRRH 752
| | | | |
DB 241 AVFRLMEOCMAVEPGORPSFSTIYOELOSIRKRRH 275
| | | | |

RESULT 11

US-10-464-805-30
; Sequence 30, Application US/10464805
; Publication No. US20040005648A1
; GENERAL INFORMATION:
; APPLICANT: LEV, SIMA
; APPLICANT: SCHLESSINGER, JOSEPH
; TITLE OF INVENTION: PYK2 RELATED PRODUCTS AND METHODS
; FILE REFERENCE: 038602/0116
; CURRENT APPLICATION NUMBER: US/10/464,805
; PRIOR FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: US/09/165,062
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 08/460,626
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: 08/357,642
; PRIOR FILING DATE: 1994-12-15
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-464-805-30

Query Match 23.5%; Score 911; DB 15; Length 251;

Best Local Similarity 67.1%; Pred. No. 1.1e-46;

Matches 169; Conservative 36; Mismatches 45; Indels 2; Gaps 2;

QY 493 LGEQIGRGNGFGEVFSGRRLADNTLVAVKSCRETLPPDLKAKFLQEARILKOYSHPIVRL 552
| | | | |
DB 2 LGEILGKGNFGFVYKGTLLK-DKTPVAVKTCCKEDLPOLKIKFLQEARILKOYDHPIVRL 60
| | | | |
QY 553 IGVCTOKOPITYIYVMEIVOGGDFLTFLRTGRLRVKTLQWVGAAGMEYLESKCIIHR 612
| | | | |
DB 61 IGVCTOKOPITYIYVMEIVOGGDFLTFLRTGRLRVKTLQWVGAAGMEYLESKCIIHR 120
| | | | |
QY 613 DLAAARNCVTEKYNLKIISDFGMSREADVAAAGGRLROVPVKTAPALNAYGRYSSED 672
| | | | |
DB 121 DLAAARNCVTEKYNLKIISDFGMSREADVAAAGGRLROVPVKTAPALNAYGRYSSED 179
| | | | |
QY 673 VMSFGILMETFSIGASPYPLNSNOQTRFEVEKGRLLPCPELCPDVAFLMEOCMAVEPG 732
| | | | |
DB 180 VMSFGILMETFSIGASPYPLNSNOQTRFEVEKGRLLPCPELCPDVAFLMEOCMAVEPG 239
| | | | |
QY 733 ORPSFSTIYOE 744
| | | | |
DB 240 NRPKFSDLHKEL 251
| | | | |

RESULT 12

US-08-987-689A-30
; Sequence 30, Application US/08987689A
; Publication No. US20020048782A1
; GENERAL INFORMATION:
; APPLICANT: Sima Lev

APPLICANT: Joseph Schlessinger
TITLE OF INVENTION: PYK2 RELATED PRODUCTS AND METHODS
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/987,689A

FILING DATE: December 9, 1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/032,824

FILING DATE: December 11, 1996

APPLICATION NUMBER: 08/460,626

FILING DATE: June 2, 1995

APPLICATION NUMBER: 08/357,642

FILING DATE: December 15, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 230/110

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 251 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULAR TYPE: protein

US-08-987-689A-30

Query Match 23.2%; Score 899; DB 8; Length 251;

Best Local Similarity 66.7%; Pred. No. 5.8e-46;

Matches 168; Conservative 36; Mismatches 46; Indels 2; Gaps 2;

QY 493 LGEQIGRGNGFGEVFSGRRLADNTLVAVKSCRETLPPDLKAKFLQEARILKOYSHPIVRL 552
| | | | |
DB 2 LGEILGKGNFGFVYKGTLLK-DKTPVAVKTCCKEDLPOLKIKFLQEARILKOYDHPIVRL 60
| | | | |
QY 553 IGVCTOKOPITYIYVMEIVOGGDFLTFLRTGRLRVKTLQWVGAAGMEYLESKCIIHR 612
| | | | |
DB 61 IGVCTOKOPITYIYVMEIVOGGDFLTFLRTGRLRVKTLQWVGAAGMEYLESKCIIHR 120
| | | | |
QY 613 DLAAARNCVTEKYNLKIISDFGMSREADVAAAGGRLROVPVKTAPALNAYGRYSSED 672
| | | | |
DB 121 DLAAARNCVTEKYNLKIISDFGMSREADVAAAGGRLROVPVKTAPALNAYGRYSSED 179
| | | | |
QY 673 VMSFGILMETFSIGASPYPLNSNOQTRFEVEKGRLLPCPELCPDVAFLMEOCMAVEPG 732
| | | | |
DB 180 VMSFGILMETFSIGASPYPLNSNOQTRFEVEKGRLLPCPELCPDVAFLMEOCMAVEPG 239
| | | | |
QY 733 ORPSFSTIYOE 744
| | | | |
DB 240 NRPKFSDLHKEL 251
| | | | |

RESULT 13

US-10-292-524-30
; Sequence 30, Application US/10292524
; Publication No. US20030119067A1
; GENERAL INFORMATION:


```

; Sequence 42, Application US/09863776
; Publication No. US20030198953A1
; GENERAL INFORMATION:
; APPLICANT: Szytek, Kimberly A
; APPLICANT: Majumder, Kumud
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vitehu
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spaderne, Steven K
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Li, Li
; APPLICANT: Taupier, Raymond J
; APPLICANT: Gansolli, Esra
; TITLE OF INVENTION: No. US20030198953A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-020
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/206,679
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 60/206,688
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 60/206,829
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 60/207,748
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 60/207,798
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 60/208,263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 60/208,831
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/209,451
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: 60/210,060
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: 60/219,507
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: 60/221,337
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/221,927
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 60/263,135
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,688
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/263,694
; PRIOR FILING DATE: 2001-01-24
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 42
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Tyrosine
; OTHER INFORMATION: Kinase Consensus Sequence
US-09-863-776-42

```

```

Query Match 17.2%; Score 668.5; DB 10; Length 257;
Best Local Similarity 49.0%; Pred. No. 3.4e-32;
Matches 127; Conservative 53; Mismatches 72; Indels 7; Gaps 4;

QY 491 LVIGSGIGNGRGEVSSGR---ADNTIVAVKSCRETLPPDLKAKFLQEARILKQYSHP 547
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 LTLGKLLGSGAGFGEVYKGLTKGKGKEVEVAATLKEDASEQIIEFLREAKIMRDLDP 60

QY 548 NIVRLIGVCTOKOPIYIYVIELVQGDPLFTFLRTGAR-LRVKTLIQWVGDDAAAGMEYLES 606
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 NIVKLGVCTEEPRLMIVMEYMEGDLDTLRKRRKELSLSDLSFALQIARGMEYLES 120

```

```

QY 607 KCCIRHDLAARACLYTEKRVLKISDFGMSRE-EADGVVAASGGLRQVPYKMTAPREALNYG 665
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 KNFVHRDLAARACLYGENKTVKIDPGLSRDLYSDIYVYKGG--KLPTRMAAPESLKEG 178

QY 666 RYSSSESDVMSFGIILMETFSLGASFPYPLNSQOTREPVKGGRLPCPELCPDAVFLRMEQ 725
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 179 KFTSKSDVMSFGVLLMEIFTLGESPYPGMSNEVEYLKGYRLPQPPNCPPBIYDLMQ 238

QY 726 CWAYEPGQRPSPSTTYQEL 744
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 239 CWADPEDERPTESELVERL 257

```

Search completed: March 18, 2005, 16:12:24
Job time : 148 secs

This Page Blank (uspto)

Db 361 QEALQGLQVALCSQAQLQAQOELLQTKLEHLGPEBPPLLLQDDRHSTSSSEOREGGR 420
QY 421 TPTLEILKSHISGIRFPKPSNLYRLEGEGPSPILPIDHLSTQOPLTKKSGVLAHRAVP 480
Db 421 TPTLEILKSHISGIRFPKPSNLYRLEGEGPSPILPIDHLSTQOPLTKKSGVLAHRAVP 480
QY 481 KKKWVNLHEDLVLEGOIGRGNFGEVFSGRRLADNTLVAAVSCRETIPLPKAKFLOEARI 540
Db 481 KKKWVNLHEDLVLEGOIGRGNFGEVFSGRRLADNTLVAAVSCRETIPLPKAKFLOEARI 540
QY 541 LKOYSHPNIVRLIGVCTOKOPIYIMELVQGGDFLFTLFTEGARLVKTLQWVGDAAG 600
Db 541 LKOYSHPNIVRLIGVCTOKOPIYIMELVQGGDFLFTLFTEGARLVKTLQWVGDAAG 600
QY 601 MEYLESKCCIHRLDIAARNCLVTEKNVCLKISDFGMSREBAGVYAASGGLRQVPVKWTABE 660
Db 601 MEYLESKCCIHRLDIAARNCLVTEKNVCLKISDFGMSREBAGVYAASGGLRQVPVKWTABE 660
QY 661 ALNYGRYSSESDVWSFGILLMETFSLGASPPYNLNQOTREVEYKCGRLPCBELCPDAVF 720
Db 661 ALNYGRYSSESDVWSFGILLMETFSLGASPPYNLNQOTREVEYKCGRLPCBELCPDAVF 720
QY 721 RLMEQCMAYEPGQRPSPSTIYQELQSIKRRHR 752
Db 721 RLMEQCMAYEPGQRPSPSTIYQELQSIKRRHR 752

RESULT 2
US-10-003-295-2
Sequence 2, Application US/10003295
Patent No. 6686187
GENERAL INFORMATION:
APPLICANT: GAN, Weiniu et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: CL001183DIV
CURRENT APPLICATION NUMBER: US/10/003,295
CURRENT FILING DATE: 2001-12-06
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 752
TYPE: PR1
ORGANISM: Homo sapiens
US-10-003-295-2

Query Match 100.0%; Score 3879; DB 4; Length 752;
Best Local Similarity 100.0%; Pred. No. 1.1e-261;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGFSSSELCSPOGHGVLQOQOEAEFLRLLEGKRMKMAORVVSDBRYAGLLHHMSIQDSGGQS 60
Db 1 MGFSSSELCSPOGHGVLQOQOEAEFLRLLEGKRMKMAORVVSDBRYAGLLHHMSIQDSGGQS 60
QY 61 RAISPSPISSQSWAEITTSQTEGSLRLRQHAEDLNSGPKSLSLIRERQOLKRTYSEOM 120
Db 61 RAISPSPISSQSWAEITTSQTEGSLRLRQHAEDLNSGPKSLSLIRERQOLKRTYSEOM 120
QY 121 QOIQOELTTHSODIEKLKSQYRALARDAQAQKRYQASKDKORDKADKYVRSIMKLF 180
Db 121 QOIQOELTTHSODIEKLKSQYRALARDAQAQKRYQASKDKORDKADKYVRSIMKLF 180
QY 181 AHNHNRVVLGVRAAQLHGHQHLPLGLRLSLQDLHEBMACILKETLOEYLETSSLYODE 240
Db 181 AHNHNRVVLGVRAAQLHGHQHLPLGLRLSLQDLHEBMACILKETLOEYLETSSLYODE 240
QY 241 VVAIHHEMAAARARQPEAEYQGLROYGSAPVPCVTFDESILBEGPLEBGLQLANE 300
Db 241 VVAIHHEMAAARARQPEAEYQGLROYGSAPVPCVTFDESILBEGPLEBGLQLANE 300
QY 301 LTVESVOHTLTSVTDELAVAATEMVFROEMVTOLOQELRNEBENTHPRERVOLLGKROYL 360
Db 301 LTVESVOHTLTSVTDELAVAATEMVFROEMVTOLOQELRNEBENTHPRERVOLLGKROYL 360

Db 301 LTVESVOHTLTSVTDELAVAATEMVFROEMVTOLOQELRNEBENTHPRERVOLLGKROYL 360
QY 361 QEALQGLQVALCSQAQLQAQOELLQTKLEHLGPEBPPLLLQDDRHSTSSSEOREGGR 420
Db 361 QEALQGLQVALCSQAQLQAQOELLQTKLEHLGPEBPPLLLQDDRHSTSSSEOREGGR 420
QY 421 TPTLEILKSHISGIRFPKPSNLYRLEGEGPSPILPIDHLSTQOPLTKKSGVLAHRAVP 480
Db 421 TPTLEILKSHISGIRFPKPSNLYRLEGEGPSPILPIDHLSTQOPLTKKSGVLAHRAVP 480
QY 481 KKKWVNLHEDLVLEGOIGRGNFGEVFSGRRLADNTLVAAVSCRETIPLPKAKFLOEARI 540
Db 481 KKKWVNLHEDLVLEGOIGRGNFGEVFSGRRLADNTLVAAVSCRETIPLPKAKFLOEARI 540
QY 541 LKOYSHPNIVRLIGVCTOKOPIYIMELVQGGDFLFTLFTEGARLVKTLQWVGDAAG 600
Db 541 LKOYSHPNIVRLIGVCTOKOPIYIMELVQGGDFLFTLFTEGARLVKTLQWVGDAAG 600
QY 601 MEYLESKCCIHRLDIAARNCLVTEKNVCLKISDFGMSREBAGVYAASGGLRQVPVKWTABE 660
Db 601 MEYLESKCCIHRLDIAARNCLVTEKNVCLKISDFGMSREBAGVYAASGGLRQVPVKWTABE 660
QY 661 ALNYGRYSSESDVWSFGILLMETFSLGASPPYNLNQOTREVEYKCGRLPCBELCPDAVF 720
Db 661 ALNYGRYSSESDVWSFGILLMETFSLGASPPYNLNQOTREVEYKCGRLPCBELCPDAVF 720
QY 721 RLMEQCMAYEPGQRPSPSTIYQELQSIKRRHR 752
Db 721 RLMEQCMAYEPGQRPSPSTIYQELQSIKRRHR 752

RESULT 3
US-09-817-180-4
Sequence 4, Application US/09817180
Patent No. 6340584
GENERAL INFORMATION:
APPLICANT: GAN, Weiniu et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: CL001183
CURRENT APPLICATION NUMBER: US/09/817,180
CURRENT FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 822
TYPE: PR1
ORGANISM: Human
US-09-817-180-4

Query Match 98.8%; Score 3814; DB 3; Length 822;
Best Local Similarity 91.5%; Pred. No. 1.6e-258;
Matches 752; Conservative 0; Mismatches 0; Indels 70; Gaps 1;

QY 1 MGFSSSELCSPOGHGVLQOQOEAEFLRLLEGKRMKMAORVVSDBRYAGLLHHMSIQDSGGQS 60
Db 1 MGFSSSELCSPOGHGVLQOQOEAEFLRLLEGKRMKMAORVVSDBRYAGLLHHMSIQDSGGQS 60
QY 61 RAISPSPISSQSWAEITTSQTEGSLRLRQHAEDLNSGPKSLSLIRERQOLKRTYSEOM 120
Db 61 RAISPSPISSQSWAEITTSQTEGSLRLRQHAEDLNSGPKSLSLIRERQOLKRTYSEOM 120
QY 121 QOIQOELTTHSODIEKLKSQYRALARDAQAQKRYQASKDKORDKADKYVRSIMKLF 180
Db 121 QOIQOELTTHSODIEKLKSQYRALARDAQAQKRYQASKDKORDKADKYVRSIMKLF 180
QY 181 AHNHNRVVLGVRAAQLHGHQHLPLGLRLSLQDLHEBMACILKETLOEYLETSSLYODE 240
Db 181 AHNHNRVVLGVRAAQLHGHQHLPLGLRLSLQDLHEBMACILKETLOEYLETSSLYODE 240
QY 241 VVAIHHEMAAARARQPEAEYQGLROYGSAPVPCVTFDESILBEGPLEBGLQLANE 300
Db 241 VVAIHHEMAAARARQPEAEYQGLROYGSAPVPCVTFDESILBEGPLEBGLQLANE 300

Db 241 VVAIHREMAAAAR1QPEAEYQGFRLQYGSADVPVPCVTFDESLIEGEPLEBGEQLQJLNE 300
QY 301 LTVESVQHTLTSVTDELAVATEMVFRRQEMVTLQOQELRNEBENTHPRERVOLLGKROYL 360
Db 301 LTVESVQHTLTSVTDELAVATEMVFRRQEMVTLQOQELRNEBENTHPRERVOLLGKROYL 360
QY 361 QBALQGLQVALCSQAKLQAOQELQTKLEHLGEPGPPVLLIQDDRHSSTSSBOEREGGR 420
Db 361 QBALQGLQVALCSQAKLQAOQELQTKLEHLGEPGPPVLLIQDDRHSSTSSBOEREGGR 420
QY 421 TPTLEILKSHISGIFRPKFS----- 440
Db 421 TPTLEILKSHISGIFRPKFSLEPPQLQIPVOKPLHEQLMWHGALPRAEVAELVHSGDF 480
QY 441 -----NLRYLEGBGPPSIPLLIDHLSTQOPLTKK 470
Db 481 LVRESQKQEVYLVLMQGLPRHFIIQSLDNLRYLEGBGPPSIPLLIDHLSTQOPLTKK 540
QY 471 SGVVLHRAVPKDKWVLNHEDLVLGEOIGRNGEVPFSGRLRADNTLVAVKSCRETLPPDL 530
Db 541 SGVVLHRAVPKDKWVLNHEDLVLGEOIGRNGEVPFSGRLRADNTLVAVKSCRETLPPDL 600
QY 531 KAKPLQEARILKQYSHPNIVRLIGVCTOKQPIYIWMELVQGGDFLFLTEGARLRVKTL 590
Db 601 KAKPLQEARILKQYSHPNIVRLIGVCTOKQPIYIWMELVQGGDFLFLTEGARLRVKTL 660
QY 591 LQWGDAAAGMEYLSKCCIHRLDLARNCLVTEKNVLKISDFGMSHEADGVYAASGGLR 650
Db 661 LQWGDAAAGMEYLSKCCIHRLDLARNCLVTEKNVLKISDFGMSHEADGVYAASGGLR 720
QY 651 QVPVKMTABEALNYGRYSSESDVWSFGILLMETFSIGASPYNLSNQOTREFVEKGRPL 710
Db 721 QVPVKMTABEALNYGRYSSESDVWSFGILLMETFSIGASPYNLSNQOTREFVEKGRPL 780
QY 711 CPELCDAVFRLEMCMAYPEGQRPSPSTIYQELQSIKRRHR 752
Db 781 CPELCDAVFRLEMCMAYPEGQRPSPSTIYQELQSIKRRHR 822

RESULT 4
US-10-003-295-4
; Sequence 4, Application US/10003295
; Patent No. 6686187
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01183DIV
; CURRENT APPLICATION NUMBER: US/10/003,295
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-003-295-4

Query Match 98.8%; Score 3834; DB 4; Length 822;
Best Local Similarity 91.5%; Pred. No. 1.6e-258;
Matches 752; Conservative 0; Mismatches 0; Indels 70; Gaps 1;

QY 1 MGFSSSELCPQGHGVLQOMQEAELRLLEGKRMMAQVYSDREYAGLLHMSLQDSGGOS 60
Db 1 MGFSSSELCPQGHGVLQOMQEAELRLLEGKRMMAQVYSDREYAGLLHMSLQDSGGOS 60
QY 61 RAISPSPISQSWAETTSQTGSLRLRQHAEDLNSGPIKSLILRERQQLKTYSEOW 120
Db 61 RAISPSPISQSWAETTSQTGSLRLRQHAEDLNSGPIKSLILRERQQLKTYSEOW 120
QY 121 QQLQOGLTHTHSQDIETKLSQYALARDSAQAQKRYQASKXDRKAKDKYVRSIMKLF 180

Db 121 QQLQOGLTHTHSQDIETKLSQYALARDSAQAQKRYQASKXDRKAKDKYVRSIMKLF 180
QY 181 AHHNRYVLCVRAAOIHHQHHOILLPGLLRSLQODLHEBACTIKELQETYLEISLVODE 240
Db 181 AHHNRYVLCVRAAOIHHQHHOILLPGLLRSLQODLHEBACTIKELQETYLEISLVODE 240
QY 241 VVAIHREMAAAAR1QPEAEYQGFRLQYGSADVPVPCVTFDESLIEGEPLEBGEQLQJLNE 300
Db 241 VVAIHREMAAAAR1QPEAEYQGFRLQYGSADVPVPCVTFDESLIEGEPLEBGEQLQJLNE 300
QY 301 LTVESVQHTLTSVTDELAVATEMVFRRQEMVTLQOQELRNEBENTHPRERVOLLGKROYL 360
Db 301 LTVESVQHTLTSVTDELAVATEMVFRRQEMVTLQOQELRNEBENTHPRERVOLLGKROYL 360
QY 361 QBALQGLQVALCSQAKLQAOQELQTKLEHLGEPGPPVLLIQDDRHSSTSSBOEREGGR 420
Db 361 QBALQGLQVALCSQAKLQAOQELQTKLEHLGEPGPPVLLIQDDRHSSTSSBOEREGGR 420
QY 421 TPTLEILKSHISGIFRPKFS----- 440
Db 421 TPTLEILKSHISGIFRPKFSLEPPQLQIPVOKPLHEQLMWHGALPRAEVAELVHSGDF 480
QY 441 -----NLRYLEGBGPPSIPLLIDHLSTQOPLTKK 470
Db 481 LVRESQKQEVYLVLMQGLPRHFIIQSLDNLRYLEGBGPPSIPLLIDHLSTQOPLTKK 540
QY 471 SGVVLHRAVPKDKWVLNHEDLVLGEOIGRNGEVPFSGRLRADNTLVAVKSCRETLPPDL 530
Db 541 SGVVLHRAVPKDKWVLNHEDLVLGEOIGRNGEVPFSGRLRADNTLVAVKSCRETLPPDL 600
QY 531 KAKPLQEARILKQYSHPNIVRLIGVCTOKQPIYIWMELVQGGDFLFLTEGARLRVKTL 590
Db 601 KAKPLQEARILKQYSHPNIVRLIGVCTOKQPIYIWMELVQGGDFLFLTEGARLRVKTL 660
QY 591 LQWGDAAAGMEYLSKCCIHRLDLARNCLVTEKNVLKISDFGMSHEADGVYAASGGLR 650
Db 661 LQWGDAAAGMEYLSKCCIHRLDLARNCLVTEKNVLKISDFGMSHEADGVYAASGGLR 720
QY 651 QVPVKMTABEALNYGRYSSESDVWSFGILLMETFSIGASPYNLSNQOTREFVEKGRPL 710
Db 721 QVPVKMTABEALNYGRYSSESDVWSFGILLMETFSIGASPYNLSNQOTREFVEKGRPL 780
QY 711 CPELCDAVFRLEMCMAYPEGQRPSPSTIYQELQSIKRRHR 752
Db 781 CPELCDAVFRLEMCMAYPEGQRPSPSTIYQELQSIKRRHR 822

RESULT 5
US-09-538-092-866
; Sequence 866, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Iolc
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIOR APPLICATION NUMBER: 2000-03-29
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 866
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P07332
US-09-538-092-866

Query Match 98.7%; Score 3828; DB 4; Length 822;
 Best Local Similarity 91.4%; Pred. No. 4.2e-258;
 Matches 751; Conservative 0; Mismatches 1; Indels 70; Gaps 1;

```

QY 1 MGFSSSLCPQGHGVLQMOQEAELRLLEGKMKMAQRYKSDREYAGLLHMSLQDSGGGS 60
DB 1 MGFSSSLCPQGHGVLQMOQEAELRLLEGKMKMAQRYKSDREYAGLLHMSLQDSGGGS 60
QY 61 RAISPSPIQSMAETTSQTGSLRLRQHAEDLNSGPISKSLIRERQOLKRTYSEOW 120
DB 61 RAISPSPIQSMAETTSQTGSLRLRQHAEDLNSGPISKSLIRERQOLKRTYSEOW 120
QY 121 QOLQOELTKTHSODIKLSQYRALARDSAOKRYQEA-KDORDKADKYRSLMKLF 180
DB 121 QOLQOELTKTHSODIKLSQYRALARDSAOKRYQEA-KDORDKADKYRSLMKLF 180
QY 181 AHNHRYVLGVRAAQLHNNHNNHQLLPGLRLSLQDLHEMACILKEILQEYLEISLVQDE 240
DB 181 AHNHRYVLGVRAAQLHNNHNNHQLLPGLRLSLQDLHEMACILKEILQEYLEISLVQDE 240
QY 241 VVAIHREMAAARLQPEAEYQGLFQYGSAPDVPCTFDESLBEGEPLEBGELOQNE 300
DB 241 VVAIHREMAAARLQPEAEYQGLFQYGSAPDVPCTFDESLBEGEPLEBGELOQNE 300
QY 301 LTVESYQHTLTSYTDLAVALATEMVFRRQEMVTQLOQELRNEBENTHPRERVOLIGROYL 360
DB 301 LTVESYQHTLTSYTDLAVALATEMVFRRQEMVTQLOQELRNEBENTHPRERVOLIGROYL 360
QY 361 QEALQGLQVALCSQAQLQAOQELLQTLKHLGEBPPVLLQDDRHSTSSBOEREGR 420
DB 361 QEALQGLQVALCSQAQLQAOQELLQTLKHLGEBPPVLLQDDRHSTSSBOEREGR 420
QY 421 TPRLTLKSHISGIFRPKFS-----NLVRLGEGEPSPILIDHLLSTQOPLTKK 440
DB 421 TPRLTLKSHISGIFRPKFS-----NLVRLGEGEPSPILIDHLLSTQOPLTKK 440
QY 441 -----NLVRLGEGEPSPILIDHLLSTQOPLTKK 470
DB 441 -----NLVRLGEGEPSPILIDHLLSTQOPLTKK 470
QY 481 LVRESQKQEVYLVLMQGLPRFIITQSLDNLVRLGEGEPSPILIDHLLSTQOPLTKK 540
DB 481 LVRESQKQEVYLVLMQGLPRFIITQSLDNLVRLGEGEPSPILIDHLLSTQOPLTKK 540
QY 541 SGVTLHRAVPKDKMVLNHEDLVLGEOIGRNGFGEVSGRLRADNTLVAVKSCRETLPPDL 600
DB 541 SGVTLHRAVPKDKMVLNHEDLVLGEOIGRNGFGEVSGRLRADNTLVAVKSCRETLPPDL 600
QY 601 KAKFLQEARILKQYSHPNIVRLIGVCTQKQPIYIMVELVQGGDFLTFRTGRLRVKTL 660
DB 601 KAKFLQEARILKQYSHPNIVRLIGVCTQKQPIYIMVELVQGGDFLTFRTGRLRVKTL 660
QY 651 LQWVGDAAGMEYLEBCKCIHRDLAARNCLVTEKNVLKISDFGMSREBAGVYAAAGSLR 650
DB 651 LQWVGDAAGMEYLEBCKCIHRDLAARNCLVTEKNVLKISDFGMSREBAGVYAAAGSLR 650
QY 710 QVAVKMTAPBALNYGRYSSESDVMSFGILLMETFSLGASPYPLNSNOQTRFVPEKGRLLP 710
DB 710 QVAVKMTAPBALNYGRYSSESDVMSFGILLMETFSLGASPYPLNSNOQTRFVPEKGRLLP 710
QY 752 CPELCPDAVFRMLEQCMAYEFGORBSFTTYOGLSIRKRRH 752
DB 752 CPELCPDAVFRMLEQCMAYEFGORBSFTTYOGLSIRKRRH 752
QY 780 CPELCPDAVFRMLEQCMAYEFGORBSFTTYOGLSIRKRRH 822
DB 780 CPELCPDAVFRMLEQCMAYEFGORBSFTTYOGLSIRKRRH 822

```

RESULT 6
 US-09-886-319A-64
 ; Sequence 64, Application US/09886319A
 ; Patent No. 6586185
 ; GENERAL INFORMATION:
 ; APPLICANT: Wolf, Eckard
 ; APPLICANT: Werner, Sabine
 ; APPLICANT: Halle, Jörn-Peter
 ; APPLICANT: Regenbogen, Johannes
 ; APPLICANT: Goppelt, Andreas
 ; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for

TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
 ; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
 ; FILE REFERENCE: 50125/014002
 ; CURRENT APPLICATION NUMBER: US/09/886,319A
 ; CURRENT FILING DATE: 2001-06-20
 ; PRIOR APPLICATION NUMBER: US 60/222,081
 ; PRIOR FILING DATE: 2000-08-01
 ; PRIOR APPLICATION NUMBER: DE 10030149.5
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: FaelsEQ for Windows Version 4.0
 ; SEQ ID NO 64
 ; LENGTH: 822
 ; TYPE: prt
 ; ORGANISM: Homo sapiens
 ; US-09-886-319A-64

Query Match 47.8%; Score 1853; DB 4; Length 822;
 Best Local Similarity 45.5%; Pred. No. 1e-120;
 Matches 376; Conservative 150; Mismatches 217; Indels 84; Gaps 9;

```

QY 1 MGFSSSLCPQGHGVLQMOQEAELRLLEGKMKMAQRYKSDREYAGLLHMSLQDSGGGS 60
DB 1 MGFSSDL--KNSHEAVLKLQDWELRLLETVKKFMALRIKSDREYASTQNLQNDKEST 58
QY 61 RAISPSPIQSMAETTSQTGSLRLRQHAEDLNSGPISKSLIRERQOLKRTYSEOW 120
DB 59 VQMNIVSVNSKSMMLMIQTEQLSRIMKTHADLNSGPHRLTMIKQYKKSIGYGH 118
QY 121 QOLQOELTKTHSODIKLSQYRALARDSAOKRYQEA-SKORDKADKYRSLMKLF 179
DB 119 QOIBAKMTKVTELEKLCYRQIKEMNSAKERYKALAGKETBKAKERYDATMTKL 178
QY 180 FAHNHRYVLGVRAAQLHNNHNNHQLLPGLRLSLQDLHEMACILKEILQEYLEISLVQD 239
DB 179 HMLHQQYVALAKGAOLHQQYVDITPLPLDLSLQMOEBMIRALNGIPDEYQSLTSLYE 238
QY 240 BVVAIHREMAAARLQPEAEYQGLFQYGSAPDVPCTFDESLBEGEPLEBGELOQNE 299
DB 239 ETVANHKEIQMSVEQIDPSTETNNFTDVHRTTAAKEQIEFTSLSEBENENLOANEIMYN 298
QY 300 ELTVESYQHTLTSYTDLAVALATEMVFRRQEMVTQLOQELRNEBENTHPR-ERYVOLIGKRO 358
DB 299 NLTAESLQWMLKTLABELMQOQMLNKEAVLBEKRIEESSECEKKSIDVILLSQRO 358
QY 359 VIQELQGLQVALCSQAQLQAOQELLQTLKHLGEBPPVLLQDDRHSTSSBOEREGR 418
DB 359 ALBELKQSVQOLRCCTEAKFSQKELLEQVQENDGKEPPVNVYEDDASVTSMERKE-- 416
QY 419 GRTPLTLKSHISGIFR-PK-----NLVRLGEGEPSPILIDHLLSTQ 438
DB 417 -RLSKPESTRHSIAGLINSFKSAVSSSALSDMISISEKRLAODMYHAIIPREAQELKL 475
QY 439 -----FSLVRLGEGEPSPILIDHLLSTQ 464
DB 476 KQGDPLVESHGKPEBYLVSVSDGQRHFIITQYDVMYRFGGTGFSNIPQIIDHLYTTX 535
QY 465 QPLTKKSGVTLHRAVPKDKMVLNHEDLVLGEOIGRNGFGEVSGRLRADNTLVAVKSCR 523
DB 536 QVITKKSQGVVLTNPDKQKMTLSHEDVLTGELKKGNGFGEVYKGLK-DKTSVAVKTKK 594
QY 524 ETLPPDLAKFLQEARILKQYSHPNIVRLIGVCTQKQPIYIMVELVQGGDFLTFRTGGA 583
DB 595 EDLPQELKTLQEAQKILKQYDHPNIVKLIQVCTQKQPIYIMVELVSGGDFLTFRRKKD 654
QY 584 RLKVTLLQWVGDAAGMEYLEBCKCIHRDLAARNCLVTEKNVLKISDFGMSREBAGVY 643
DB 655 ELKQLQYKVSJDAAGMLYLESKNCIHRDLAARNCLVGENNVKLKISDFGMSRQEDGGYV 714
QY 644 AASGGLQVAVKMTAPBALNYGRYSSESDVMSFGILLMETFSLGASPYPLNSNOQTRFV 703
DB 715 SSS-GIKQIPKMTAPBALNYGRYSSESDVMSFGILLMETFSLGVCYPGDMTQOAREOV 773

```


[illegible]

RESULT 9
 US-08-701-191A-38
 Sequence 38, Application US/08701191A
 Patent No. 5942428
 GENERAL INFORMATION:
 APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
 APPLICANT: and Steven R. Hubbard
 TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
 OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Filth Street
 STREET: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071-2066
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 MEDIUM TYPE: storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: FastISO for Windows 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/701,191A
 FILING DATE: August 21, 1996
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Walburg, Richard J.
 REGISTRATION NUMBER: 32,327
 REFERENCE/DOCKET NUMBER: 227/088

```

? TELECOMMUNICATION INFORMATION
? TELEPHONE: (213) 489-1600
? TELEFAX: (213) 955-0440
?
? TELEEX: 67-3510
?
? INFORMATION FOR SEQ ID NO: 38:
?
? SEQUENCE CHARACTERISTICS:
? LENGTH: 275 amino acids
? TYPE: amino acid
? STANDARDS: single
? TOPOLOGY: linear
?
? MOLECULE TYPE: protein
?
? OS-08-701-191A-38

```

Query Match	37.2%	Score 1444;	DB 2;	Length 275;
Best Local Similarity	99.6%	Pred. No. 7.4e-93;		
Matches 274; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

Qy	478	AVPKDQKVLNHNHDLVLGEOIGRGNGEYFSRLADNTLVAVKSCRETLPPDLKAKFLOE	537
Db	1	AVPKDKVLNHNHDLVLGEOIGRGNGEYFSRLADNTLVAVKSCRETLPPDLKAKFLOE	60
Qy	538	ARILKQYSHPNIVALLIGVCTOKOPYIYIMELVOGDFLPLRTGABLRKYTLLOWGDA	597
Db	61	ARILKQYSHPNIVRLIGVCTOKOPYIYIMELVOGDFLPLRTGABLRKYTLLOWGDA	120
Qy	598	AAGMEYLEKSCCIHNDLAARNCVLYEKVNLXISDPGMSREADSVYAASGGLRQVPKWT	657
Db	121	AAGMEYLEKSCCIHNDLAARNCVLYEKVNLXISDPGMSREADSVYAASGGSQVPKWT	180
Qy	658	APBALNYGYSSESQVWVSFGILLMETFSLGASPPYNNLSNOOTREFVEKGRLLPCBELCPD	717
Db	181	APBALNYGYSSESQVWVSFGILLMETFSLGASPPYNNLSNOOTREFVEKGRLLPCBELCPD	240
Qy	718	AVFRLMEQCMAYEPQORSFSTIYVELOSIKRHR	752
Db	241	AVFRLMEQCMAYEPQORSFSTIYVELOSIKRHR	275

```

RESULT 10
US-09-664-526-38
: Sequence 38, Application US/09664526
: Patent No. 6682921
: GENERAL INFORMATION:
: APPLICANT: MOHAMMADI, MOOSA
: APPLICANT: SCHLESINGER, JOSEPH
: APPLICANT: HUBBARD, STEVAN R.
: TITLE OF INVENTION: RECEPTOR TYROSINE KINASES
: TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
: FILE REFERENCE: 038602/0847
: CURRENT APPLICATION NUMBER: US/09/664,526
: CURRENT FILING DATE: 2000-09-18
: PRIOR APPLICATION NUMBER: 09/188, 809
: PRIOR FILING DATE: 1998-11-09
: PRIOR APPLICATION NUMBER: 08/701,191
: PRIOR FILING DATE: 1996-08-21
: NUMBER OF SEQ ID NOS: 41
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 38
: LENGTH: 275
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-664-526-38

```

Qy	478	AVPDKQWVLNHHDLVLGEQIGRGNGFGEVSGRLRADNTLVAVKSCRETLPPDLKAKFLQOE	537
Db	1	AVPDKQWVLNHHEDVLGBQIGRGNGFGEVSGRLRADNTLVAVKSCRETLPPDLKAKFLQOE	60
Qy	538	ARLLKQSHNPVAVLLIGVCTQKQETIYVMEVLQGGDFLFLRTBEARLRVKTLLQWVGDA	597

Db 61 ARLKQVSHPNIVRLIGVCTQKQPIYIWMELVGGDFLFLRTGARGRLVKTLQWVGDA 120
QY 588 AAGMEYLSKCCIRHDLAARNCLVTEKNVLKISDFGMSREADGVYAASGGLRQVPVKMT 657
Db 121 AAGMEYLSKCCIRHDLAARNCLVTEKNVLKISDFGMSREADGVYAASGGSROVPVKMT 180
QY 658 APEALNYGYSSESDVWSFGILLMETFSLGASPPYPLNSNOQTRFVEKGRLLPCPELCPD 717
Db 181 APEALNYGYSSESDVWSFGILLMETFSLGASPPYPLNSNOQTRFVEKGRLLPCPELCPD 240
QY 718 AVFRLMEQCMAYEPGQRPSPSTIYQELQSIKRRH 752
Db 241 AVFRLMEQCMAYEPGQRPSPSTIYQELQSIKRRH 275

RESULT 11

US-07-857-224B-62
/ Sequence 62, Application US/07857224B
/ Patent No. 5958784
/ GENERAL INFORMATION:
/ APPLICANT: Benner, Steven A.
/ TITLE OF INVENTION: Predicting Folded Structures of Proteins
/ NUMBER OF SEQUENCES: 114
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Steven A. Benner
/ STREET: Hadlaubstrasse 151
/ CITY: Zurich
/ STATE: none
/ COUNTRY: Switzerland
/ ZIP: (note: this is an international post code) CH-8092
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
/ OPERATING SYSTEM: Macintosh 7.0
/ SOFTWARE: Microsoft Word
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/857,224B
/ FILING DATE: 03/25/92
/ CLASSIFICATION: 436
/ PRIOR APPLICATION DATA: none
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (international) 41 1 632 2830
/ TELEFAX: (international) 41 1 262 2437
/ TELEX: none
/ INFORMATION FOR SEQ ID NO: 62:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 262
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE:
/ DESCRIPTION: protein
/ ORIGINAL SOURCE:
/ ORGANISM: human
/ FEATURE: Protein kinase; Table 8 Column 70
/ PUBLICATION INFORMATION:
/ AUTHORS:
/ AUTHORS: Hanke, S. K.
/ AUTHORS: Quinn, A. M.
/ AUTHORS: Hunter, T.
/ TITLE: The protein kinase family
/ JOURNAL: Science
/ VOLUME: 241
/ PAGES: 42-52
/ DATE: 1988
/ US-07-857-224B-62

Query Match 35.2%; Score 1364.5; DB 2; Length 262;
Best Local Similarity 99.6%; Pred. No. 2.4e-87;
Matches 262; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 489 EDVLVGEQIGRGNGFVSGRLADNTLVAVKSCRETLPPDLKAKFLQBARILKQYSHN 548
Db 1 EDVLVGEQIGRGNGFVSGRLADNTLVAVKSCRETLPPDLKAKFLQBARILKQYSHN 60

QY 549 IYRLIGVCTQKQPIYIWMELVGGDFLFLRTGARGRLVKTLQWVGDAAGMEYLSKC 608
Db 61 IYRLIGVCTQKQPIYIWMELVGGDFLFLRTGARGRLVKTLQWVGDAAGMEYLSKC 120
QY 609 C1HRDLAARNCLVTEKNVLKISDFGMSREADGVYAASGGLRQVPVKMTAPEALNYGRYS 668
Db 121 C1HRDLAARNCLVTEKNVLKISDFGMSREADGVYAAS-GLRQVPVKMTAPEALNYGRYS 179
QY 669 SSSDVWSFGILLMETFSLGASPPYPLNSNOQTRFVEKGRLLPCPELCPDAVFLMEQCA 728
Db 180 SSSDVWSFGILLMETFSLGASPPYPLNSNOQTRFVEKGRLLPCPELCPDAVFLMEQCA 239
QY 729 YEPGQRPSPSTIYQELQSIKRRH 751
Db 240 YEPGQRPSPSTIYQELQSIKRRH 262

RESULT 12

US-07-857-224B-63
/ Sequence 63, Application US/07857224B
/ Patent No. 5958784
/ GENERAL INFORMATION:
/ APPLICANT: Benner, Steven A.
/ TITLE OF INVENTION: Predicting Folded Structures of Proteins
/ NUMBER OF SEQUENCES: 114
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Steven A. Benner
/ STREET: Hadlaubstrasse 151
/ CITY: Zurich
/ STATE: none
/ COUNTRY: Switzerland
/ ZIP: (note: this is an international post code) CH-8092
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
/ OPERATING SYSTEM: Macintosh 7.0
/ SOFTWARE: Microsoft Word
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/857,224B
/ FILING DATE: 03/25/92
/ CLASSIFICATION: 436
/ PRIOR APPLICATION DATA: none
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (international) 41 1 632 2830
/ TELEFAX: (international) 41 1 262 2437
/ TELEX: none
/ INFORMATION FOR SEQ ID NO: 63:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 262
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE:
/ DESCRIPTION: protein
/ ORIGINAL SOURCE:
/ ORGANISM: cat
/ FEATURE: Protein kinase; Table 8 Column 71
/ PUBLICATION INFORMATION:
/ AUTHORS:
/ AUTHORS: Hanke, S. K.
/ AUTHORS: Quinn, A. M.
/ AUTHORS: Hunter, T.
/ TITLE: The protein kinase family
/ JOURNAL: Science
/ VOLUME: 241
/ PAGES: 42-52
/ DATE: 1988
/ US-07-857-224B-63

Query Match 34.8%; Score 1350.5; DB 2; Length 262;
Best Local Similarity 97.7%; Pred. No. 2.2e-86;
Matches 257; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 489 EDLVIGEIOGRNGFSEVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKQYSHN 548
DB 1 EDLVIGEIOGRNGFSEVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKQYSHN 60
QY 549 IYRLIGVCTOKOPIYIMVELVGGDFLTLRTGRLRVKTLLOWGDAAGMEYLESKC 608
DB 61 IYRLIGVCTOKOPIYIMVELVGGDFLTLRTGRLRVKTLLOWGDAAGMEYLESKC 120
QY 609 CIHRDLAARNCLVTEKNVLTISDFGMSREADVYAASGGLRQVPYKWTAPALNYGRYS 668
DB 121 CIHRDLAARNCLVTEKNVLTISDFGMSREADVYAAS-GLRQVPYKWTAPALNYGRYS 179
QY 669 SSSDVMSFGILLMETFSLGASPYPLNSNOQTRFEVKEGRLPCPELCPDAVRLMEQCA 728
DB 180 SSSDVMSFGILLMETFSLGASPYPLNSNOQTRFEVKEGRLPCPELCPDAVRLMEQCA 239
QY 729 YEPGQRPFSFTIYQELQSIKRRH 751
DB 240 YEPGQRPFSFTIYQELQSIKRRH 262

RESULT 13

US-07-857-224B-64
Sequence 64, Application US/07857224B
Patent No. 5958784

GENERAL INFORMATION:

APPLICANT: Benner, Steven A.

TITLE OF INVENTION: Predicting Folded Structures of Proteins

NUMBER OF SEQUENCES: 114

CORRESPONDENCE ADDRESS:

ADDRESSEE: Steven A. Benner

STREET: Hadlaubstrasse 151

CITY: Zurich

STATE: none

COUNTRY: Switzerland

ZIP: (note: this is an international post code) CH-8092

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage

COMPUTER: Apple Macintosh

OPERATING SYSTEM: Macintosh 7.0

SOFTWARE: Microsoft Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/857,224B

FILING DATE: 03/25/92

CLASSIFICATION: 436

PRIOR APPLICATION DATA: none

TELECOMMUNICATION INFORMATION:

TELEPHONE: (international) 41 1 632 2830

TELEFAX: (international) 41 1 262 2437

TELEX: none

INFORMATION FOR SEQ ID NO: 64:

SEQUENCE CHARACTERISTICS:

LENGTH: 262

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

DESCRIPTION: protein

ORGANISM: chicken

FEATURE: Protein kinase; Table 8 Column 72

PUBLICATION INFORMATION:

AUTHORS: Hanka, S. K.

AUTHORS: Quinn, A. M.

AUTHORS: Hunter, T.

TITLE: The protein kinase family

JOURNAL: Science

VOLUME: 241

PAGES: 42-52

DATE: 1988

US-07-857-224B-64

Query Match

28.8%; Score 1116.5; DB 2; Length 262;

Best Local Similarity 79.8%; Pred. No. 4,2e-70;
Matches 210; Conservative 29; Mismatches 23; Indels 1; Gaps 1;
QY 489 EDLVIGEIOGRNGFSEVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKQYSHN 548
DB 1 EDLVIGEIOGRNGFSEVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKQYSHN 60
QY 549 IYRLIGVCTOKOPIYIMVELVGGDFLTLRTGRLRVKTLLOWGDAAGMEYLESKC 608
DB 61 IYRLIGVCTOKOPIYIMVELVGGDFLTLRTGRLRVKTLLOWGDAAGMEYLESKC 120
QY 609 CIHRDLAARNCLVTEKNVLTISDFGMSREADVYAASGGLRQVPYKWTAPALNYGRYS 668
DB 121 CIHRDLAARNCLVTEKNVLTISDFGMSREADVYAASGGLRQVPYKWTAPALNYGRYS 179
QY 669 SSSDVMSFGILLMETFSLGASPYPLNSNOQTRFEVKEGRLPCPELCPDAVRLMEQCA 728
DB 180 SSSDVMSFGILLMETFSLGASPYPLNSNOQTRFEVKEGRLPCPELCPDAVRLMEQCA 239
QY 729 YEPGQRPFSFTIYQELQSIKRRH 751
DB 240 YEPGQRPFSFTIYQELQSIKRRH 262

RESULT 14

US-10-003-690-8
Sequence 8, Application US/10003690
Patent No. 6787345

GENERAL INFORMATION:

APPLICANT: Roy A.J. Curtis

TITLE OF INVENTION: 55053, A No. 6787345e1 Human Eukaryotic Kinase

TITLE OF INVENTION: and Uses Therefor

FILE REFERENCE: MNI-206

CURRENT APPLICATION NUMBER: US/10/003,690

PRIOR FILING DATE: 2001-11-15

PRIOR APPLICATION NUMBER: 60/248,893

NUMBER OF SEQ ID NOS: 8

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 8

LENGTH: 280

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: tyrkin_6 domain

US-10-003-690-8

Query Match 16.6%; Score 642; DB 4; Length 280;

Best Local Similarity 46.3%; Pred. No. 4.7e-37;

Matches 133; Conservative 49; Mismatches 65; Indels 40; Gaps 8;

QY 491 IYRLIGVCTOKOPIYIMVELVGGDFLTLRTGRLRVKTLLOWGDAAGMEYLESKC 608
DB 1 IYRLIGVCTOKOPIYIMVELVGGDFLTLRTGRLRVKTLLOWGDAAGMEYLESKC 120
QY 549 IYRLIGVCTOKOPIYIMVELVGGDFLTLRTGRLRVKTLLOWGDAAGMEYLESKC 608
DB 61 IYRLIGVCTOKOPIYIMVELVGGDFLTLRTGRLRVKTLLOWGDAAGMEYLESKC 120
QY 609 CIHRDLAARNCLVTEKNVLTISDFGMSREADVYAASGGLRQVPYKWTAPALNYGRYS 668
DB 121 CIHRDLAARNCLVTEKNVLTISDFGMSREADVYAASGGLRQVPYKWTAPALNYGRYS 179
QY 669 SSSDVMSFGILLMETFSLGASPYPLNSNOQTRFEVKEGRLPCPELCPDAVRLMEQCA 728
DB 180 SSSDVMSFGILLMETFSLGASPYPLNSNOQTRFEVKEGRLPCPELCPDAVRLMEQCA 239
QY 729 YEPGQRPFSFTIYQELQSIKRRH 751
DB 240 YEPGQRPFSFTIYQELQSIKRRH 262

RESULT 15

QY 706 GGLRCPB-----LCPDAVRLMEQCAVPEPGQRPFSFTIYQEL 744
DB 234 GYRLKPENDLPISSVTCPELDLYDLMLQCMADPEBDRPFFSELVERL 280

```
US-09-949-016-6798
; Sequence 6798, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 6798
; LENGTH: 1148
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6798

Query Match      16.1%; Score 624.5; DB 4; Length 1148;
Best Local Similarity 36.0%; Pred. No. 4.5e-35;
Matches 138; Conservative 68; Mismatches 128; Indels 49; Gaps 10;

QY 405 DNHSTSSSEQRERGGRTPTLEILKSHISGIFRPKFS----- 440
DB 141 EKHSMYHGVSNAAEYF-----LSSGINSFLVRESESSPSQGISLRYEGRVYHRYNT 196
QY 441 ----NLYRLEGGFPSIPLLIH-----LIST--QQPLTKSGVVLHRAVPK-DKVTLN 487
DB 197 ASDGKLVYSSSRFNTLAEVLVHHSTVADGLITTLHYPAKRNKPTVGVSPNYDKEME 256
QY 488 HEDVLVBGIGRGNGEVSFSGRLADNTLVAVKSCRE-TLPDYLAKKFLQEARILKQYSH 546
DB 257 RDTITMKHKLGGQGEVYEGVWKYSLTVAVKTKEDTWEVE--EFLKEAIVMKELRH 313
QY 547 PNIIVRLIGVCTOKOPIYIYIMELVQGGDFLPLRTGEGARLRVK--TLLOMVGDAAGMEYL 604
DB 314 PVLVQLLGVCYTRPPYIITTEFTYGNLDYLR-ECNROEVNAVVLVMAIQISSAMEYL 372
QY 605 ESKCCIHRDLAARNCLVTEKNVLIKISDFGMSRBEADGVYASGGLRQVPVKWTAPBALNY 664
DB 373 EKKNFIRHDLAARNCLVGENHLLVKVADFGLSRLMTGDTYTAHAGAK-PIKWTAPESLAY 431
QY 665 GRYSSESQVWVSGILLMETFSLGASPYPNLSNQOTREFVEKGRLLPCPELCPDAVFRIME 724
DB 432 NKFISKSDVMAFGVLLMEIATYGMSPYPOIDRSQVYELLEKDYRMKRFBGCEKYEELMR 491
QY 725 QCMAYEPGQPSFSTIYQELQSI 747
DB 492 ACMQWNPSPDRPSFAELHQAFTM 514
```

Search completed: March 18, 2005, 16:01:08
Job time : 57 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 15:50:13 ; Search time 45 Seconds

(without alignments)
1607.889 Million cell updates/sec

Title: US-10-660-763-2

Perfect score: 3879

Sequence: 1 MGFSSELCSPOGHGVLOQMG.....QRPSTTYQELGSLRKHR 752

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 79: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3834	98.8	822	1	TVHUF protein-tyrosine k
2	3620	93.3	820	1	TVCTPF protein-tyrosine k
3	3473	89.5	820	2	148347 protein-tyrosine k
4	2635	67.9	824	2	150618 c-fos proto onco
5	2570	66.3	873	1	TVFVFS protein-tyrosine k
6	2561	66.0	873	1	TVFVF protein-tyrosine k
7	2479	63.9	609	1	TVMWGC protein-tyrosine k
8	1980	51.0	477	1	TVMWCS protein-tyrosine k
9	1853	47.8	822	1	TVHUF protein-tyrosine k
10	1450	37.4	533	1	TVFVFP protein-tyrosine k
11	1160	29.9	803	1	OKPFS protein-tyrosine k
12	1099.5	28.3	323	2	S04328 protein-tyrosine k
13	1097	28.3	453	2	149663 protein-tyrosine k
14	734	18.9	181	2	150406 tyrosine kinase (f
15	632	16.3	697	2	A26132 proto-fps protein
16	623	16.1	1123	2	A39662 gag-abl-pol polyp
17	622	16.0	1130	1	TVHUA kinase-related tra
18	620.5	16.0	981	1	FOVMGM protein-tyrosine k
19	619.5	16.0	334	2	S24552 gag-abl polypotei
20	608	15.7	505	2	159296 protein-tyrosine k
21	605.5	15.6	1146	2	B35962 protein-tyrosine k
22	605.5	15.6	1182	2	A35962 protein-tyrosine k
23	604	15.6	465	2	149826 protein-tyrosine k
24	603.5	15.6	527	2	149133 protein-tyrosine k
25	603.5	15.6	1196	2	T23832 protein-tyrosine k
26	603.5	15.6	1520	1	TVFPA protein-tyrosine k
27	602	15.5	507	2	A55625 protein-tyrosine k
28	600.5	15.5	467	2	156579 protein-tyrosine k
29	597	15.4	527	2	S13763 protein-tyrosine k

30	595.5	15.4	362	2	S24551 protein-tyrosine k
31	592.5	15.3	621	2	15697 protein-tyrosine k
32	591	15.2	527	2	184883 protein-tyrosine k
33	589	15.2	675	2	S60612 protein-tyrosine k
34	588.5	15.2	620	1	S33253 protein-tyrosine k
35	586.5	15.1	1028	2	A45388 protein-tyrosine k
36	583	15.0	507	1	A3939 protein-tyrosine k
37	583	15.0	2340	2	148310 kinase-related pro
38	581.5	15.0	680	4	138491 nucleophosmin/anap
39	581	15.0	630	1	T01380 protein-tyrosine k
40	580.5	15.0	496	2	A56040 protein-tyrosine k
41	579.5	14.9	402	1	TVFVUR protein-tyrosine k
42	579.5	14.9	505	1	S24550 protein-tyrosine k
43	579	14.9	2347	1	TVHURS kinase-related pro
44	578.5	14.9	590	1	TVFVFS protein-tyrosine k
45	578	14.9	1068	2	151670 focal adhesion kin

ALIGNMENTS

RESULT 1

TVHUF

protein-tyrosine kinase (EC 2.7.1.112) fes/fps - human

Query Match	98.8%	Score 3834;	DB 1;	Length 822;
Best Local Similarity	91.5%	Pred. No. 7.8e-141;		
Matches 752; Conservative	0;	Mismatches 0;	Indels 70;	Gaps 1;
QY	1	MGFSSELCSPOGHGVLOQMGAEALRLBGRKMAQRVVSDDREYAGLHHMSLSDSGGOS	60	
DB	1	MGFSSELCSPOGHGVLOQMGAEALRLBGRKMAQRVVSDDREYAGLHHMSLSDSGGOS	60	
QY	61	RAISPDSPISQSWAETTSQTBGLSRLLRQHAADLNSGPKSLSLIREQOQLRTKYSBOW	120	
DB	61	RAISPDSPISQSWAETTSQTBGLSRLLRQHAADLNSGPKSLSLIREQOQLRTKYSBOW	120	
QY	121	QQLQGLTFTTHSQDBLKLKSYRRLARSDAQAQRKRYQASXDKDNDKADKTVRSLSMLKF	180	

```

Db      121 OQLQGLTHTSHSDIEKLSQRYALARDSDQAQKRYQESKQDROKADKXVRSIMKLF 180
Qy      181 AHNHRYVLGVRAAQLHHQHHLHLLPGLRLSLDLDHEMACTIKELIQEYLSISLVODE 240
Db      181 AHNHRYVLGVRAAQLHHQHHLHLLPGLRLSLDLDHEMACTIKELIQEYLSISLVODE 240
Qy      241 VVAIHREMAAAARIQPEAEYQGFRLQYGSAPVPCVTFDESLLEGEPLPBGELQANE 300
Db      241 VVAIHREMAAAARIQPEAEYQGFRLQYGSAPVPCVTFDESLLEGEPLPBGELQANE 300
Qy      301 LTVESVQHTLTSVTDELAVATEMVFRRQEWVTLQOGLRNEBENTHPRERVOILLGRQYL 360
Db      301 LTVESVQHTLTSVTDELAVATEMVFRRQEWVTLQOGLRNEBENTHPRERVOILLGRQYL 360
Qy      361 QEALQGLQVALCSQAQLQAOQELLQTKLEHLGEGEPPLVLLDODRHSTSSSEQEREGR 420
Db      361 QEALQGLQVALCSQAQLQAOQELLQTKLEHLGEGEPPLVLLDODRHSTSSSEQEREGR 420
Qy      421 TPTLEILKSHISGIFRPKFS----- 440
Db      421 TPTLEILKSHISGIFRPKFS----- 440
Qy      441 -----NLYRLEGESEPSIPLLDLHLLSTQOPLTKK 470
Db      481 LVRESQKQEVYLVLMQGLPRHFIIQSIDNLYRLEGESEPSIPLLDLHLLSTQOPLTKK 540
Qy      471 SGVVLHRAVPKQKWLNLHEDLVLEQIQRGNFGEVPSGRLRADNTLVAVKSCRETLPPDL 530
Db      541 SGVVLHRAVPKQKWLNLHEDLVLEQIQRGNFGEVPSGRLRADNTLVAVKSCRETLPPDL 600
Qy      531 KAKFLQEARILKQYSHPNIVRLIGVCTQKQPIYIWMELVQGGDFLTLFTTEGARLRVKTL 590
Db      601 KAKFLQEARILKQYSHPNIVRLIGVCTQKQPIYIWMELVQGGDFLTLFTTEGARLRVKTL 660
Qy      591 LQWVGDAAGMEYLESKCCIHRLDLAARNCILVTEKNVLIKISDFGMSREADGVYAASGGLR 650
Db      661 LQWVGDAAGMEYLESKCCIHRLDLAARNCILVTEKNVLIKISDFGMSREADGVYAASGGLR 720
Qy      651 QVPVKMTAPEALNYGRYSSESQVMSFGILLMETFSIGASPYPNLSNQOTREFVEKGRLLP 710
Db      721 QVPVKMTAPEALNYGRYSSESQVMSFGILLMETFSIGASPYPNLSNQOTREFVEKGRLLP 780
Qy      711 CPELCDAVFRLEMCQWAVEPQGRPSFSTIYQELQSIKRRHR 752
Db      781 CPELCDAVFRLEMCQWAVEPQGRPSFSTIYQELQSIKRRHR 822

```

RESULT 2

TVCTPF

protein-tyrosine kinase (BC 2.7.1.112) fes/fps - cat

C/Species: Felis silvestris catus (domestic cat)

C/Date: 30-Sep-1989 #sequence_revision 30-Jun-1992 #text_change 23-Feb-1997

C/Accession: A27824

R./Roebroek, A.J.M.; Schalken, J.A.; Omekink, C.; Bloemers, H.P.J.; Van de Ven, W.J.M.

J. Virol. 61, 2009-2016, 1987

A./Title: Structure of the feline c-fes/fps proto-oncogene: geneis of a retroviral onco

A./Reference number: A27824; PMID:87198954; PMID:3553615

A./Accession: A27824

A./Molecule type: DNA

A./Residues: 1-820 <ROB>

C/Genetics:

A./Gene: fes/fps

C./Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology

C./Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; proto-oncogene

F/458-543/Domain: SH2 homology <SH2>

F/557-819/Domain: protein kinase homology <KIN>

F/565-573/Region: protein kinase ATP-binding motif

F/588/Active site: Lys #status predicted

F/711/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 93.3%; Score 3620; DB 1; Length 820;

Best Local Similarity 85.8%; Pred. No. 1.3e-132;

```

Matches 705; Conservative 21; Mismatches 24; Indels 72; Gaps 2;
Qy      1 MGFSEELSPQGHGVLQMOAEALRLLEGMRKMAQRYVSDSEYVGLLHHMSLDQSGGS 60
Db      1 MGFSEELSPQGHGVLQMOAEALRLLEGMRKMAQRYVSDSEYVGLLHHMSLDQSGG-- 58
Qy      61 RAISPDSPISQSWAEITQTEGLSRLLROHAEIDLNSGPKSLILIRERQURKTYSEGM 120
Db      59 RGTGYSPISQSWAEITQTEGLSRLLROHAEIDLNSGPKSLILIRERQURKTYSEGM 118
Qy      121 OQLQGLTHTSHSDIEKLSQRYALARDSDQAQKRYQESKQDROKADKXVRSIMKLF 180
Db      119 OQLQGLTHTSHSDIEKLSQRYALARDSDQAQKRYQESKQDROKADKXVRSIMKLF 178
Qy      181 AHNHRYVLGVRAAQLHHQHHLHLLPGLRLSLDLDHEMACTIKELIQEYLSISLVODE 240
Db      179 AHNHRYVLGVRAAQLHHQHHLHLLPGLRLSLDLDHEMACTIKELIQEYLSISLVODE 238
Qy      241 VVAIHREMAAAARIQPEAEYQGFRLQYGSAPVPCVTFDESLLEGEPLPBGELQANE 300
Db      239 VVAIHREMAAAARIQPEAEYQGFRLQYGSAPVPCVTFDESLLEGEPLPBGELQANE 298
Qy      301 LTVESVQHTLTSVTDELAVATEMVFRRQEWVTLQOGLRNEBENTHPRERVOILLGRQYL 360
Db      299 LTVESVQHTLTSVTDELAVATEMVFRRQEWVTLQOGLRNEBENTHPRERVOILLGRQYL 358
Qy      361 QEALQGLQVALCSQAQLQAOQELLQTKLEHLGEGEPPLVLLDODRHSTSSSEQEREGR 420
Db      359 QEALQGLQVALCSQAQLQAOQELLQTKLEHLGEGEPPLVLLDODRHSTSSSEQEREGR 418
Qy      421 TPTLEILKSHISGIFRPKFS----- 440
Db      419 TPTLEILKSHISGIFRPKFS----- 440
Qy      441 -----NLYRLEGESEPSIPLLDLHLLSTQOPLTKK 470
Db      479 LVRESQKQEVYLVLMQQPRHFIIQSADNLYRLEGESEPSIPLLDLHLLSTQOPLTKK 538
Qy      471 SGVVLHRAVPKQKWLNLHEDLVLEQIQRGNFGEVPSGRLRADNTLVAVKSCRETLPPDL 530
Db      539 SGVVLHRAVPKQKWLNLHEDLVLEQIQRGNFGEVPSGRLRADNTLVAVKSCRETLPPDL 598
Qy      531 KAKFLQEARILKQYSHPNIVRLIGVCTQKQPIYIWMELVQGGDFLTLFTTEGARLRVKTL 590
Db      599 KAKFLQEARILKQYSHPNIVRLIGVCTQKQPIYIWMELVQGGDFLTLFTTEGARLRVKTL 658
Qy      591 LQWVGDAAGMEYLESKCCIHRLDLAARNCILVTEKNVLIKISDFGMSREADGVYAASGGLR 650
Db      659 LQWVGDAAGMEYLESKCCIHRLDLAARNCILVTEKNVLIKISDFGMSREADGVYAASGGLR 718
Qy      651 QVPVKMTAPEALNYGRYSSESQVMSFGILLMETFSIGASPYPNLSNQOTREFVEKGRLLP 710
Db      719 QVPVKMTAPEALNYGRYSSESQVMSFGILLMETFSIGASPYPNLSNQOTREFVEKGRLLP 778
Qy      711 CPELCDAVFRLEMCQWAVEPQGRPSFSTIYQELQSIKRRHR 752
Db      779 CPELCDAVFRLEMCQWAVEPQGRPSFSTIYQELQSIKRRHR 820

```

RESULT 3

I48347

protein-tyrosine kinase-related protein - mouse

N./Alternate names: c-fes protein

C./Species: Mus musculus (house mouse)

C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004

C/Accession: I48347; J01112

R./Wilks, A.F.; Kurban, R.R.

Oncogene 3, 289-294, 1988

A./Title: Isolation and structural analysis of murine c-fes cDNA clones.

A./Reference number: I48347; PMID:89083198; PMID:3060793

A./Accession: I48347

A./Status: preliminary; translated from GB/EMBL/DBJ

A./Molecule type: mRNA

A:Residues: 1-820 <RES>
A:Cross-references: UNIPROT:P16879; EMBL:X12616; NID:g50955; PIDD:CAA1138.1; PID:g50956
R:Wills, A.F.; Kurban, R.R.; Hovens, C.M.; Ralph, S.U.
Gene 85, 67-74, 1989
A:Title: The application of the polymerase chain reaction to cloning members of the pro
A:Reference number: JH0112; MUID:90152381; PMID:2482828
A:Accession: JH0112
A:Molecule type: mRNA
A:Residues: 678-745 <WIL>
A:Experimental source: haemopoietic cell, clone FD15
C:Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C:Keywords: ATP
F:458-543/Domain: SH2 homology <SH2>
F:557-819/Domain: protein kinase homology <KIN>
F:565-573/Region: protein kinase ATP-binding motif

Query Match 89.5%; Score 3473; DB 2; Length 820;
Best Local Similarity 82.4%; Pred. No. 66-127;
Matches 677; Conservative 35; Mismatches 38; Indels 72; Gaps 2;

QY 1 MGFSSSLCPQGHGVLQOQOEALRLLEGKRMMAQRYVSDREYAGLHHMSIQDSGGQS 60
DB 1 MGFSSSLCPQGHGVLQOQOEALRLLEGKRMMAQRYVSDREYAGLHHMSIQDSGGQS 60
QY 61 RAISPSPIQSQAETTSOTEGSLRLRQHAEDLNSGPKSLRLRBOQLRKYTSBQM 120
DB 61 WSGSPSPSQSAETTSOTEGSLRLRQHAEDLNSGPKSLRLRBOQLRKYTSBQM 120
QY 121 QOLQOELTHTSHODIEKLKSQYRALARDSAQAKRYQASKDKDRKADKYVRSIMKLF 180
DB 121 QOLQOELTHTSHODIEKLKSQYRALARDSAQAKRYQASKDKDRKADKYVRSIMKLF 180
QY 181 AHNHRYVLGVRAQLHQNHHQQLLPGLRLSLQDLHEBMACTIKELQYELTSSIVQD 240
DB 181 AHNHRYVLGVRAQLHQNHHQQLLPGLRLSLQDLHEBMACTIKELQYELTSSIVQD 240
QY 241 VVAIHHEMAAARIQPEAEVQGLRQYGSAPRVPVCVTFDESILREGELEPGELOLMB 300
DB 241 VVAIHHEMAAARIQPEAEVQGLRQYGSAPRVPVCVTFDESILREGELEPGELOLMB 300
QY 301 LTVESVQHTLTSTVDLAVATEMVFRRQEMVTLQOELRNEBENTHPRERVOVLGKQYL 360
DB 301 LTVESVQHTLTSTVDLAVATEMVFRRQEMVTLQOELRNEBENTHPRERVOVLGKQYL 360
QY 361 QEALQGLQVALCSQAQKLAQOQELLQTKLEHLGEPPEPVLLQDDRHSTSSSEQEEGR 420
DB 361 QEALQGLQVALCSQAQKLAQOQELLQTKLEHLGEPPEPVLLQDDRHSTSSSEQEEGR 420
QY 421 TPTLEILKSHISGIFRPKFS----- 440
DB 421 TPTLEILKSHISGIFRPKFS----- 440
QY 441 -----NLVRLGEGFPSPILPLIDHLSTQOPLTKK 470
DB 441 -----NLVRLGEGFPSPILPLIDHLSTQOPLTKK 470
QY 479 LVRESQKQEVYLVVMMDGHRHFIIOQLDNLVRLGEGDFSPILPLIDHLSTQOPLTKK 538
DB 479 LVRESQKQEVYLVVMMDGHRHFIIOQLDNLVRLGEGDFSPILPLIDHLSTQOPLTKK 538
QY 471 SGVILHRAVPKDKWLVNHEDLVLGEOIQRNGFGEVFGRLRADNTLVAVKSCRETLPPDL 530
DB 471 SGVILHRAVPKDKWLVNHEDLVLGEOIQRNGFGEVFGRLRADNTLVAVKSCRETLPPDL 530
QY 531 KAKPELQEAIIKQYSHPNIVRLIGVCTQOPRIYIMVELVOGGDFLTFRLTEGARLKVKT 590
DB 531 KAKPELQEAIIKQYSHPNIVRLIGVCTQOPRIYIMVELVOGGDFLTFRLTEGARLKVKT 590
QY 591 LQWVGDAAGMEYLESKCCHRDLAARNCLVTEKNVLTISDFGMSREEDAGVVAASGLR 650
DB 591 LQWVGDAAGMEYLESKCCHRDLAARNCLVTEKNVLTISDFGMSREEDAGVVAASGLR 650
QY 651 LQWVGDAAGMEYLESKCCHRDLAARNCLVTEKNVLTISDFGMSREEDAGVVAASGLR 718
DB 651 LQWVGDAAGMEYLESKCCHRDLAARNCLVTEKNVLTISDFGMSREEDAGVVAASGLR 718
QY 719 QVAVKKTATAPBALNYGYSSSDVMSFGILLMETFSLGASDPYPLNSQQTREFEYKGRLLP 778
DB 719 QVAVKKTATAPBALNYGYSSSDVMSFGILLMETFSLGASDPYPLNSQQTREFEYKGRLLP 778
QY 711 CPBLCPDAVFRMLMEQCMAYEPQGRPSFTIYQELQIRKRR 752

DB 779 CPBLCPDAVFRMLMEQCMAYEPQGRPSFTIYQELQIRKRR 820

RESULT 4
c-fps proto oncogene - chicken
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence, revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: 150618
R:Huang, C.C.; Hammond, C.; Bishop, J.M.
J. Mol. Biol. 181, 175-186, 1985
A:Title: Nucleotide sequence and topography of chicken c-fps. Geneis of a retroviral onc
A:Reference number: 150618; MUID:85160839; PMID:3879969
A:Accession: 150618
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-824 <HUA>
A:Cross-references: UNIPROT:Q90778; EMBL:X02266; NID:g63203; PIDD:CAA26155.1; PID:g87104;
C:Genetics:
A:Gene: c-fps
A:Intons: 75/3; 133/3; 166/1; 227/2; 273/2; 313/2; 354/2; 416/3; 442/3; 553/3; 5;
C:Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C:Keywords: ATP
F:462-547/Domain: SH2 homology <SH2>
F:561-823/Domain: protein kinase homology <KIN>
F:569-577/Region: protein kinase ATP-binding motif

Query Match 67.9%; Score 2635; DB 2; Length 824;
Best Local Similarity 62.7%; Pred. No. 1,1e-94;
Matches 518; Conservative 99; Mismatches 133; Indels 76; Gaps 3;

QY 1 MGFSSSLCPQGHGVLQOQOEALRLLEGKRMMAQRYVSDREYAGLHHMSIQDS 56
DB 1 MGFSSSLCPQGHGVLQOQOEALRLLEGKRMMAQRYVSDREYAGLHHMSIQDS 56
QY 57 GGQSRALSDSPISQMAITTSOTEGSLRLRQHAEDLNSGPKSLRLRBOQLRKY 116
DB 57 GGQSRALSDSPISQMAITTSOTEGSLRLRQHAEDLNSGPKSLRLRBOQLRKY 116
QY 121 SEWQOQLOELTHTSHODIEKLKSQYRALARDSAQAKRYQASKDKDRKADKYVRSIMKLF 176
DB 121 SEWQOQLOELTHTSHODIEKLKSQYRALARDSAQAKRYQASKDKDRKADKYVRSIMKLF 176
QY 177 WKLFANHRYVLGVRAQLHQNHHQQLLPGLRLSLQDLHEBMACTIKELQYELTSSIVQD 236
DB 177 WKLFANHRYVLGVRAQLHQNHHQQLLPGLRLSLQDLHEBMACTIKELQYELTSSIVQD 236
QY 237 VQDEVAIIRBMAAARIQPEAEVQGLRQYGSAPRVPVCVTFDESILREGELEPGELOLMB 296
DB 237 VQDEVAIIRBMAAARIQPEAEVQGLRQYGSAPRVPVCVTFDESILREGELEPGELOLMB 296
QY 297 QNLTLVESVQHTLTSTVDLAVATEMVFRRQEMVTLQOELRNEBENTHPRERVOVLGKQYL 356
DB 297 QNLTLVESVQHTLTSTVDLAVATEMVFRRQEMVTLQOELRNEBENTHPRERVOVLGKQYL 356
QY 301 QNLTLVESVQHTLTSTVDLAVATEMVFRRQEMVTLQOELRNEBENTHPRERVOVLGKQYL 360
DB 301 QNLTLVESVQHTLTSTVDLAVATEMVFRRQEMVTLQOELRNEBENTHPRERVOVLGKQYL 360
QY 357 RQVLEALQGLQVALCSQAQKLAQOQELLQTKLEHLGEPPEPVLLQDDRHSTSSSEQEEGR 416
DB 357 RQVLEALQGLQVALCSQAQKLAQOQELLQTKLEHLGEPPEPVLLQDDRHSTSSSEQEEGR 416
QY 417 EGGRPTLEILKSHISGIFRPKFS----- 440
DB 417 EGGRPTLEILKSHISGIFRPKFS----- 440
QY 441 -----NLVRLGEGFPSPILPLIDHLSTQOPLTKK 478
DB 441 -----NLVRLGEGFPSPILPLIDHLSTQOPLTKK 478
QY 479 LVRESQKQEVYLVVMMDGHRHFIIOQLDNLVRLGEGDFSPILPLIDHLSTQOPLTKK 538
DB 479 LVRESQKQEVYLVVMMDGHRHFIIOQLDNLVRLGEGDFSPILPLIDHLSTQOPLTKK 538
QY 467 LTKSGVILHRAVPKDKWLVNHEDLVLGEOIQRNGFGEVFGRLRADNTLVAVKSCRETL 526
DB 467 LTKSGVILHRAVPKDKWLVNHEDLVLGEOIQRNGFGEVFGRLRADNTLVAVKSCRETL 526
QY 539 ITRKSGIVLTRAVALKDKWLVNHEDLVLGEOIQRNGFGEVFGRLRADNTLVAVKSCRETL 598
DB 539 ITRKSGIVLTRAVALKDKWLVNHEDLVLGEOIQRNGFGEVFGRLRADNTLVAVKSCRETL 598
QY 527 PRDLKAKPELQEAIIKQYSHPNIVRLIGVCTQOPRIYIMVELVOGGDFLTFRLTEGARL 586

```

Db      599 PPELAKFLQEARILKQYHNPNIVRLIGVCTQKQPIYIVMELVQGGDFLSFLRSEGPRLK 658
Qy      587 VKTLQMGVDAAGMYELESKCIHRDLAARNCLVTEKVKLISDPGMSREEDGYTAAS 646
Db      659 MKELIKMENAAGMEYELSKICHRDLAARNCLVTEKVKLISDPGMSREEDGYTAAS 718
Qy      647 GGLRQVPVKMTAPALNAGYSESDVMSFGILLMETFSIGASPYPNLSNOQTRFEVKG 706
Db      719 GGMKQIPVKMTAPALNAGYSESDVMSFGILLMETFSIGASPYPNLSNOQTRFEVKG 778
Qy      707 GRLPCBELCPDAVFLMEQCMAYEPGQRPSTTIYQELQSIKRRHR 752
Db      779 VRLPEPQCEPDVYRLMQRCEYDPRRRPFGAVHODLIAIRKRHR 824

RESULT 5
TVFVPS
protein-tyrosine kinase (BC 2.7.1.112) fps (clone ts) - Fujinami sarcoma virus
C:Species: Fujinami sarcoma virus
C:date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C:Accession: A26898
R:Chen, L.H.; Hacada, E.; Wheatley, W.; Lee, W.H.
Virology 155, 106-119, 1986
A:title: Single amino acid substitution, from Glu-1025 to Asp, of the fps oncogenic prot
A:Reference number: A26898; MUID:87044080; PMID:2877522
A:Accession: A26898
A:Molecule type: DNA
A:Residues: 1-873 <CH>
A:Cross-references: UNIPROT:P00530; GB:M14930; NID:G209668; PIDD:AAA42403.1; PID:G209689
C:Comment: This protein is synthesized as a gag-fps polyprotein.
C:Genetic:
A:Gene: fps
A:Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C:Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; tran
F:511-596/Domain: SH2 homology <SH2>
F:510-872/Domain: protein kinase homology <KIN>
F:618-626/Region: protein kinase ATP-binding motif
F:641/Active site: Lys #status predicted

Query Match      66.3%; Score 2570; DB 1; Length 873;
Best Local Similarity 61.4%; Pred. No. 3, 5e-92;
Matches 507; Conservative 106; Mismatches 137; Indels 76; Gaps 3;

Qy      1 MGFSSELCSPOGHVLCQMGAEELRLLEGKMMARVKSDBREYAGLHNM-----SLDS 56
Db      50 MGGPELWCPKSGSELRLQDSLRLELMKKMMSQRAASDRYAGLHNMESQLEKQSG 109
Qy      57 GGSRAISPDSPIQSMAEITQTEGLSRLLRQHAEDLNSGPLSKSLIRERQQLRKY 116
Db      110 LGHLRAITHSSQIGESWVLAQGTETLSQTLRHAEELAAAGPLAKSLIRKQQLRKY 169
Qy      117 SEWQOQLQELTTHSODITKLSQYRALARDAQAQKRYQASXKQDKDRAKQYVRS 176
Db      170 SEWQOQLQSEYAWTQOEVEKLAQYRSLVRDSTQAKRYQASXKQDKREKAKYVRS 229
Qy      177 WKLFANHRVYLGVRAAOLHNNHOLLPGILRSLODHEMACTIKELIQLYEYLSIS 236
Db      230 SKLYALHNQYVLAOVAALHNNHYORALPTLHESLYLSQOEWVLAQKELIGESCITSL 289
Qy      237 VQDEVAALHREMAAARLOPEAYOGFLRQYASADVPCCVTFDSLEEGEPLEPGEL 296
Db      290 VQGDVLAHQEVAAVEMIDPATEYSSFYQCHRYDSVPVAPYFDDSLLEBASLPGEL 349
Qy      297 QLNELTVESVQHTLTVTDELAVATEMVFRRQEMVTQLOQELFNEEBNTHPRRVOLLG 356
Db      350 QLNELTVESVQHTLTVTDELAVATEMVFRRQEMVTQLOQELFNEEBNTHPRRVOLLG 409
Qy      357 RQVLEALQGLQVLAQSAQKLOQELQTKLHLGSGPPVLLLODDHSTSSSEOR 416
Db      410 RQVLEALQGLQVLAQSAQKLOQELQTKLHLGSGPPVLLLODDHSTSSSEOR 469
Qy      417 EGGRTEPLTEIKSHISGIFRPKFS----- 440

```

```

Db      470 SG--VYALKTIKSHISGIFRPSFSLPPVPLIPEVQKPLQOAWYGAIRSEVQELTKY 527
Qy      441 -----NLYREGGSPSILLIHLSTQGP 466
Db      528 SGDFLVRSQGOEYVLSVMDGQPRHFIIOADNLYRLLEDGLPTILLIHLQSQRP 587
Qy      467 LTKSGVVLHRAVPPDKVLANHEDVLAGEQIRGNFGEVFSGRRLRADNTLVAKSCREYL 526
Db      588 ITRKSGIVLTVRAVLQDKVLANHEDVLAGEIRGNFGEVFSGRRLRADNTLVAKSCREYL 647
Qy      527 PDLAKFLQEARILKQYHNPNIVRLIGVCTQKQPIYIVMELVQGGDFLPLRTSGARLR 586
Db      648 PPELAKFLQEARILKQYHNPNIVRLIGVCTQKQPIYIVMELVQGGDFLSFLRSEGPRLK 707
Qy      587 VKTLQMGVDAAGMYELESKCIHRDLAARNCLVTEKVKLISDPGMSREEDGYTAAS 646
Db      708 MKELIKMENAAGMEYELSKICHRDLAARNCLVTEKVKLISDPGMSREEDGYTAAS 767
Qy      647 GGLRQVPVKMTAPALNAGYSESDVMSFGILLMETFSIGASPYPNLSNOQTRFEVKG 706
Db      768 GGMKQIPVKMTAPALNAGYSESDVMSFGILLMETFSIGASPYPNLSNOQTRFEVKG 827
Qy      707 GRLPCBELCPDAVFLMEQCMAYEPGQRPSTTIYQELQSIKRRHR 752
Db      828 VRLPEPQCEPDVYRLMQRCEYDPRRRPFGAVHODLIAIRKRHR 873

RESULT 6
TVFVPS
protein-tyrosine kinase (BC 2.7.1.112) fps - Fujinami sarcoma virus
C:Species: Fujinami sarcoma virus
A:Note: host Gallus gallus (chicken)
C:date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 09-Jul-2004
C:Accession: A00636
R:Shibuya, M.; Hanatua, H.
Cell 30, 787-795, 1982
A:title: Nucleotide sequence of Fujinami sarcoma virus: evolutionary relationship of its
A:Reference number: A00636; MUID:83050964; PMID:6291784
A:Accession: A00636
A:Molecule type: genomic RNA
A:Residues: 1-873 <SH1>
A:Cross-references: UNIPROT:P00530
C:Comment: This protein is synthesized as a gag-fps polyprotein.
C:Genetic:
A:Gene: fps
A:Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C:Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; poly
F:511-596/Domain: SH2 homology <SH2>
F:510-872/Domain: protein kinase homology <KIN>
F:618-626/Region: protein kinase ATP-binding motif
F:641/Active site: Lys #status predicted

Query Match      66.0%; Score 2561; DB 1; Length 873;
Best Local Similarity 61.3%; Pred. No. 7, 8e-92;
Matches 506; Conservative 106; Mismatches 138; Indels 76; Gaps 3;

Qy      1 MGFSSELCSPOGHVLCQMGAEELRLLEGKMMARVKSDBREYAGLHNM-----SLDS 56
Db      50 MGGPELWCPKSGSELRLQDSLRLELMKKMMSQRAASDRYAGLHNMESQLEKQSG 109
Qy      57 GGSRAISPDSPIQSMAEITQTEGLSRLLRQHAEDLNSGPLSKSLIRERQQLRKY 116
Db      110 LGHLRAITHSSQIGESWVLAQGTETLSQTLRHAEELAAAGPLAKSLIRKQQLRKY 169
Qy      117 SEWQOQLQELTTHSODITKLSQYRALARDAQAQKRYQASXKQDKDRAKQYVRS 176
Db      170 SEWQOQLQSEYAWTQOEVEKLAQYRSLVRDSTQAKRYQASXKQDKREKAKYVRS 229
Qy      177 WKLFANHRVYLGVRAAOLHNNHOLLPGILRSLODHEMACTIKELIQLYEYLSIS 236
Db      230 SKLYALHNQYVLAOVAALHNNHYORALPTLHESLYLSQOEWVLAQKELIGESCITSL 289

```


C:Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; tran
 F:115-200/Domain: SH2 homology <SH2>
 F:214-476/Domain: protein kinase homology <KIN>
 F:222-230/Region: protein kinase ATP-binding motif
 F:245/Active site: Lys #status predicted

Query Match 51.0%; Score 1980; DB 1; Length 477;

Best Local Similarity 80.9%; Pred. No. 9.7e-70;
 Matches 366; Conservative 11; Mismatches 10; Indels 70; Gaps 1;

```

QY 346 HPERVOLLGKROVLQALQVALCSQAQKLAQOELLQTKLHNGEPPEPPVLLQDD 405
D 1 HREGVQLAKKQVLEALQVALCSQAQKLAQOELLQTKLHNGEPPEPPVLLQDD 60
QY 406 RHSTSSSEDEREGGRPTLEILKSHISGIFRPFSLPPLQLVPEVQKPLHQLMYHGL 440
D 61 RHSTSSSEDEREGGRPTLEILKSHISGIFRPFSLPPLQLVPEVQKPLHQLMYHGL 120
QY 441 -----NLVRLGEGFPSPIL 455
D 121 PRAEVALLTHSGDFLVRESQKQEVVLSVMDGQPRHFIQADNLVPEGDFASITL 180
QY 456 LIDHLISTQOPLTKSGVLAHRAVPKQVLAHEDVLGEQIGRNGFVSGRLADNT 515
D 181 LVDFHLRSQOPLTKSGVLAHRAVPKQVLAHEDVLGEQIGRNGFVSGRLADNT 240
QY 516 LVAVVSCTEPLPDLAKFLQEARILKQVSHPIVLLIGVCTQKOPITVIMELVQGDPL 575
D 241 LVAVVSCTEPLPDLAKFLQEARILKQVSHPIVLLIGVCTQKOPITVIMELVQGDPL 300
QY 576 TELREGARLVKTLQVMDAAAGMEYLESKCIRHDLAARNCVTERKVLKISDFGMS 635
D 301 TELREGARLVKTLQVMDAAAGMEYLESKCIRHDLAARNCVTERKVLKISDFGMS 360
QY 636 REBADGVYASGGLRQVLYKMTAPBALNTRYSSSESDVMSFGILLMETFSLGASPYNLIS 695
D 361 REBADGVYASGGLRQVLYKMTAPBALNTRYSSSESDVMSFGILLMETFSLGASPYNLIS 420
QY 696 NOOTREFVKGGRLPCELPDPAVFRMLQCMAYEFGORPSFTIYOELQSIKRRR 752
D 421 NOOTREFVKGGRLPCELPDPAVFRMLQCMAYEFGORPSFTIYOELQSIKRRR 477

```

RESULT 9

TYRFP

protein-tyrosine kinase (BC 2.7.1.112) fer - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004

C:Accession: A31943; 157573

R:Haq, Q.L.; Heisterkamp, N.; Groffen, J.

Mol. Cell. Biol. 9, 1587-1593, 1989

A:Title: Isolation and sequence analysis of a novel human tyrosine kinase gene.

A:Reference number: A31943; MUID:89261786; PMID:2725517

A:Accession: A31943

A:Molecule type: mRNA

A:Residues: 1-822 <HAO>

A:Cross-references: UNIPROT:P16591; GB:J03358; NID:G339714; PIDN:AAA61190.1; PID:G339715

A:Gene: GDB:FER

A:Cross-references: GDB:125243; OMIM:176942

A:Map position: 5q12-5q14

C:Function:

A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP

C:Superfamily: protein-tyrosine kinase fgs; protein kinase homology; SH2 homology

C:Keywords: ATP; autophosphorylation; blocked amino end; 110protein; myristylation; pho

F:2-822/Product: protein-tyrosine kinase for #status predicted <MUT>

F:460-546/Domain: SH2 homology <SH2>

F:561-821/Domain: protein kinase homology <KIN>

F:569-577/Region: protein kinase ATP-binding motif

F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted

F:591/Active site: Lys #status predicted

Query Match 47.8%; Score 1853; DB 1; Length 822;

Best Local Similarity 45.5%; Pred. No. 1.3e-64;
 Matches 376; Conservative 150; Mismatches 217; Indels 84; Gaps 9;

```

QY 1 MGFSEELCSPOGHGVLQOMQEAELLRLBGMKRMQRYVSDPEYAGLLHMSLDOSGGG 60
D 1 MGFSEELCSPOGHGVLQOMQEAELLRLBGMKRMQRYVSDPEYAGLLHMSLDOSGGG 58
QY 61 RAISPDSPISQMAITSGTEGLSRLLRQHAEDLNSGFLSKSLILIREQQRKTYSEQ 120
D 59 VQMTAVSNVSKMLMIQTEQLSRIMKTHADLNSGFLRHTMTMMIKDQVYKSYIGH 118
QY 121 QQLQELTKTHSQDIEKLKSOYRALARDSAQAKRYQEA-SKDQDKAKDKRYVSLMT 179
D 119 QQLQELTKTHSQDIEKLKSOYRALARDSAQAKRYQEA-SKDQDKAKDKRYVSLMT 178
QY 180 PAHNRKRYLGVRAADLHONHQLPLGLRLDQHEMACILKEILQYELFISLVOD 239
D 179 HMLHNOYVLAALKGADLHONQYVDITPLPLLSLQKQEMIALKGIPEYQISLVLTE 238
QY 240 EVVAIHREMAAANAIOPEAEYQGFRLQYGSADVPCTPDESILLEGEPLPEGLQ 299
D 239 EIVNVHKEIQNSVEQIDSTEYNFIQVARTTAKEQIEFTSLLEENENQANEIMN 298
QY 300 ELTVESVOHTLTSVDELAVALATEVFRQENVTOLOQELRNEENTHPR-EEVQLLGK 358
D 299 NLTAESLQVMLKTLAELMOTQOMLNKEBAVLELEKRIESESCEKKSQIVLLLSQ 358
QY 359 VLQELQGLQVALCSQAQKLAQOELLQTKLHNGEPPEPPVLLQDDRHSTSSSE 418
D 359 ALEELKQSVQQLCTEAKFSQKELLEQKVOBNDKEPEPPVNVYEDARSVTSMERK 416
QY 419 GRPTLEILKSHISGIFR-PK----- 438
D 417 -RLSKFESIRHSIAGIISPKSAVSSALSMDISISEKPLAQDQWYHGAIRPEAQELL 475
QY 439 -----FSNLYRLGEGFPSPIL 464
D 476 KQGDPLVRESHGKPGSYVLSVSDGQRHFIQVDMNMYRFGTGFSPNPQILDIHYYTK 535
QY 465 OPLTKSGVLAHRAVPKQVLAHEDVLGEQIGRNGFVSGRLADNTLVAVKSR 523
D 536 QVITTKSGVLAHRAVPKQVLAHEDVLGEQIGRNGFVSGRLADNTLVAVKSR 523
QY 536 QVITTKSGVLAHRAVPKQVLAHEDVLGEQIGRNGFVSGRLADNTLVAVKSR 594
D 536 QVITTKSGVLAHRAVPKQVLAHEDVLGEQIGRNGFVSGRLADNTLVAVKSR 594
QY 524 ETLPPDLKAKFLQEARILKQVSHPIVLLIGVCTQKOPITVIMELVQGDPLTFLTEGA 583
D 524 ETLPPDLKAKFLQEARILKQVSHPIVLLIGVCTQKOPITVIMELVQGDPLTFLTEGA 583
QY 595 ETLPPDLKAKFLQEARILKQVSHPIVLLIGVCTQKOPITVIMELVQGDPLTFLTEGA 654
D 595 ETLPPDLKAKFLQEARILKQVSHPIVLLIGVCTQKOPITVIMELVQGDPLTFLTEGA 654
QY 584 RLKVTLLQVMDAAAGMEYLESKCIRHDLAARNCVTERKVLKISDFGMSREADGVY 643
D 584 RLKVTLLQVMDAAAGMEYLESKCIRHDLAARNCVTERKVLKISDFGMSREADGVY 643
QY 644 AASGGLRQVLYKMTAPBALNTRYSSSESDVMSFGILLMETFSLGASPYNLISNOQREFV 703
D 644 AASGGLRQVLYKMTAPBALNTRYSSSESDVMSFGILLMETFSLGASPYNLISNOQREFV 703
QY 715 SSS-GLKQIPIMKTAPEALNTRYSSSESDVMSFGILLMETFSLGASPYNLISNOQREFV 773
D 715 SSS-GLKQIPIMKTAPEALNTRYSSSESDVMSFGILLMETFSLGASPYNLISNOQREFV 773
QY 704 EKGGRLPCELPDPAVFRMLQCMAYEFGORPSFTIYOELQSIKRRR 750
D 774 EKGGRLPCELPDPAVFRMLQCMAYEFGORPSFTIYOELQSIKRRR 820

```

RESULT 10

TYRFP

protein-tyrosine kinase (BC 2.7.1.112) fgs - avian sarcoma virus PRCII

C:Species: avian sarcoma virus PRCII

A:Note: host Gallus gallus (chicken)

C:Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 09-Jul-2004

C:Accession: A00650

R:Huang, C.C.; Hammond, C.; Bishop, J.M.

A:Title: Nucleotide sequence of v-fps in the PRCII strain of avian sarcoma virus.

A:Reference number: A00650; MUID:84138603; PMID:6321783

A:Accession: A00650

A:Molecule type: DNA

A/Molecule type: mRNA
A/Residues: 1-323 <LET>
A/Cross-references: UNIPROT:P09760; EMBL:X13412; NID:G56169; PIDN:CAA31778.1; PID:G56170
C/Genetics:
A/Genes: Elk
C/Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C/Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; tyrosine-speci
F:62-322/Domain: protein kinase homology <KIN>
F:70-78/Region: protein kinase ATP-binding motif

Query Match 28.3%; Score 1099.5; DB 2; Length 323;
Best Local Similarity 66.2%; Pred. No. 8.9e-36;
Matches 206; Conservative 45; Mismatches 57; Indels 3; Gaps 3;

QY 441 NLYRLEGSPPSPLIDHLSTQOPLTKKSGVLLHRAVPKD-KWLNHEDVLGSGQIR 499
DB 13 NLYRLEGSPPSPLIDHLSTQOPLTKKSGVLLHRAVPKD-KWLNHEDVLGSGQIR 72
QY 500 GNGEYFSGRLRADNTLVAVKSGRETLPPDLKAKFLQEARILKQYSPNIVRLIGVCTQR 559
DB 73 GNGEYFSGRLRADNTLVAVKSGRETLPPDLKAKFLQEARILKQYSPNIVRLIGVCTQR 131
QY 560 QPIYIVMEIVQGGDELFTPLRTGEGARLVKTLQWGDAAAGMEYLSKCCIHEDLAARNC 619
DB 132 QPIYIVMEIVQGGDELFTPLRTGEGARLVKTLQWGDAAAGMEYLSKCCIHEDLAARNC 191
QY 620 LVTEKKVLTSDSGMREADGVYAASGGLRQVPVKTAPBALNYGRYSSESDVWSFGILL 679
DB 192 LVTEKKVLTSDSGMREADGVYAASGGLRQVPVKTAPBALNYGRYSSESDVWSFGILL 250
QY 680 LMEFSLGASPYPNLSNOQTRFEVFKGRLPCPELCPDAVFRLEOCMAVPEQGRPSFT 739
DB 251 LMEFSLGASPYPNLSNOQTRFEVFKGRLPCPELCPDAVFRLEOCMAVPEQGRPSFT 310
QY 740 IYQELQSIKRR 750
DB 311 LMEFSLGASPYPNLSNOQTRFEVFKGRLPCPELCPDAVFRLEOCMAVPEQGRPSFT 321

RESULT 13
149663
tyrosine kinase (fert) - mouse
C/Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Aug-2004
C/Accession: I49663
R/Fischman, K.; Edman, J.C.; Shackelford, G.M.; Turner, J.A.; Rutter, W.J.; Nir, U.
Mol. Cell. Biol. 10, 146-153, 1990
A/Title: A murine fer testis-specific transcript (fert) encodes a truncated fer protein.
A/Reference number: I49663; MUID:90097822; PMID:2294399
A/Accession: I49663
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-453 <RES>
A/Cross-references: UNIPROT:Q61561; GB:M32054; NID:G193276; PIDN:AAA37617.1; PID:G309235
C/Superfamily: protein kinase homology; SH2 homology
C/Keywords: ATP
F:91-177/Domain: SH2 homology <SH2>
F:192-452/Domain: protein kinase homology <KIN>
F:200-208/Region: protein kinase ATP-binding motif

Query Match 28.3%; Score 1097; DB 2; Length 453;
Best Local Similarity 53.3%; Pred. No. 8.6e-36;
Matches 225; Conservative 63; Mismatches 100; Indels 34; Gaps 8;

QY 348 REVYQLLGGKQVLOALQGLVALCSQAKLQAOQELL--QTKLEHLPGEPPVLLLODD 405
DB 43 KKKKRLSKFESIRHSIAGI-----IKSPKSVLGSSTVCDVIVSGRP---LAERD 90
QY 406 RHSTSSSEQ-----REGRRPTLEILKSHISGIR----KFSNLYLLEGE 448
DB 91 WYGAIPRIEAOELLKQGGFLVRESHKGEVYLVSVSDGQRHPIQFVNDLYRFEET 150
QY 449 GPFSLILLIDHLSTQOPLTKKSGVLLHRAVPKD-KWLNHEDVLGSGQIRGNGEYFVS 507

DB 151 GFSNIPOLIDHFNFKQVITKSGVLLNPIPKDKWLNHEDVLGSGQIRGNGEYFVS 210
QY 508 GRLRADNTLVAVKSGRETLPPDLKAKFLQEARILKQYSPNIVRLIGVCTQRPIYIME 567
DB 211 GRLK-DKTPVAVKTKCEDLPOLKIKFLQEARILKQYSPNIVRLIGVCTQRPIYIME 269
QY 568 LVQGGDELFTPLRTGEGARLVKTLQWGDAAAGMEYLSKCCIHEDLAARNCVTEKNVL 627
DB 270 LVQGGDELFTPLRTGEGARLVKTLQWGDAAAGMEYLSKCCIHEDLAARNCVTEKNVL 329
QY 628 KISDFGMSREADGVYAASGGLRQVPVKTAPBALNYGRYSSESDVWSFGILLMEFSLG 687
DB 330 KISDFGMSREADGVYAASGGLRQVPVKTAPBALNYGRYSSESDVWSFGILLMEFSLG 388
QY 688 ASPYPNLSNOQTRFEVFKGRLPCPELCPDAVFRLEOCMAVPEQGRPSFTIYQELQSI 747
DB 389 VCPYPMNTQOARQEVGRYMSAPQNCPEEFTIMKCMYKPNRPKFNDLHRELTVI 448
QY 748 RK 749
DB 449 RK 450

RESULT 14
150406
proto-fps protein - chicken (fragment)
C/Species: Gallus gallus (chicken)
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C/Accession: I50406
R/Pfeiff, S.L.; Zhou, R.
Virology 146, 307-314, 1995
A/Title: Defining the borders of the chicken proto-fps gene, a precursor of Fujinami sarc
A/Reference number: I50405; MUID:86020620; PMID:2996222
A/Accession: I50406
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-181 <PPA>
A/Cross-references: UNIPROT:Q90943; GB:M11611; NID:G212542; PIDN:AAA49008.1; PID:G212545
C/Genetics:
A/Insertions: 41/2, 94/1, 135/1
C/Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C/Keywords: ATP
F:1-180/Domain: protein kinase homology (fragment) <KIN>

Query Match 18.9%; Score 734; DB 2; Length 181;
Best Local Similarity 73.5%; Pred. No. 3.2e-22;
Matches 133; Conservative 26; Mismatches 22; Indels 0; Gaps 0;

QY 572 GDFLFTLRTGARLVKTLQWGDAAAGMEYLSKCCIHEDLAARNCVTEKNVLKISD 631
DB 1 GDFLFTLRTGARLVKTLQWGDAAAGMEYLSKCCIHEDLAARNCVTEKNVLKISD 60
QY 632 FQMSREADGVYAASGGLRQVPVKTAPBALNYGRYSSESDVWSFGILLMEFSLGASPY 691
DB 61 FQMSREADGVYAASGGLRQVPVKTAPBALNYGRYSSESDVWSFGILLMEFSLGASPY 120
QY 692 PULSNOQTRFEVFKGRLPCPELCPDAVFRLEOCMAVPEQGRPSFTIYQELQSIKRR 751
DB 121 PULSNOQTRFEVFKGRLPCPELCPDAVFRLEOCMAVPEQGRPSFTIYQELQSIKRR 180
QY 752 R 752
DB 181 R 181

RESULT 15
A26132
gag-ab1-pol polyprotein - feline sarcoma virus (strain Hardy-Zuckerman 2) (fragment)
N/Contains: protein-tyrosine kinase (EC 2.7.1.112)
C/Species: feline sarcoma virus
C/Date: 31-Dec-1988 #sequence_revision 09-Sep-1994 #text_change 24-Sep-1999
C/Accession: A26132

R.Bergold, P.J.; Blumenthal, J.A.; D'Andrea, E.; Snyder, H.W.; Lederman, L.; Silverstone
J. Virol. 61, 1193-1202, 1987
A>Title: Nucleic acid sequence and oncogenic properties of the HZ2 feline sarcoma virus
A.Reference number: A26132; MUID:8714138; PMID:3029415
A.Accession: A26132
A.Molecule type: DNA
A.Residues: 1-697 <BER>
A.Cross-references: GB:M15805; NID:9323863; PIDN:AAA43042.1; PID:9323864
C/Genetics:
A.Gene: gag-ab1-pol
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C/Keywords: ATP; oncogene; phosphotransferase; polypeptide; transforming protein; tyrosyl
F:69-117/Domain: SH3 homology <SH3>
F:128-218/Domain: SH2 homology <SH2>
F:241-501/Domain: protein kinase homology <KIN>
F:249-257/Region: protein kinase ATP-binding motif
F:272/Active site: Lys #status predicted

Query Match 16.3%; Score 632; DB 2; Length 697;

Best Local Similarity 32.9%; Pred.No. 1e-17; Matches 147; Indels 102; Gaps 18;

Matches 163; Conservative 84; Mismatches 147; Indels 102; Gaps 18;

QY 348 RERVOLLAGKQVLOALQGLQVALCSQAKLQAOELQTKLEHL--GPEPPVLLIQ-- 403
DB 17 RHLRLV--RQLL---LGLRGARARPTVLQVQVQCKENLLAGPSENDPNLPVALY 71
QY 404 -----DD-----RHSTSSSEQRBEGR-----TP--TLE----- 425
DB 72 DFVAGDNTLSITKGEKRLVLCYNHNGEWCEAQTKGQWPSNYITPVNSLEKHSWYHG 131
QY 426 -----ILKSHISGIRPKFS-----NLVRLGE 448
DB 132 PVSRAAEYLLSSGINSGLVRESSESPQGRSISLRYESGRVYHRYINTASDGKLYVSPES 191
QY 449 GPPSIPLLDH-----LST--QOPLTKSGVLRHRAVPK--DKWVLAHEDVLGEOIGR 499
DB 192 RNTLAEVLVHHSTVADGLITLHPAPRKPKPTVYGVSPNDKMEMERTDITMKHKLGG 251
QY 500 GNFGVFSGRLEADNTLVAVKSCRE--TLPPDLKAKFLOEARILKOYSHPNIVRLIGVCTQ 558
DB 252 GQYGEVYEGWKKVSLTVAVKTLKEDTMEVE--EFLKEAAVMKEIKHPNLVQLLGVCTR 308
QY 559 KQPIYIVMELVGGDPFLTEFLTEGARLRK--TLQWVGDAAGMEYLESKCCIHRLAA 616
DB 309 EPPFYIITEFMYGMLDYLR--ECNRQEVNAVVLTYMATQISSAMEYLEKKNFIHRDLAA 367
QY 617 RNCVTEKKNVLIKISDFGMSREAEADVAAAGGLRQVPVMTAPEALNTYGRYSSESQVMSF 676
DB 368 RNCVGENHLLVAVDFGLSLMTGDTYTHAGTK--PRIKTAPESLAVYKPSIKSDVNAF 426
QY 677 GILMETSLSGASPPYPNLSNQOTREFVEKGRLLPCBELCPDAVFRIMEQCMAYEPGQRP 736
DB 427 GVLMEIATYGMSPYPGIDLSQVYELLEKDYMERRECGCEPKYELMRACQWQNPSPDRPA 486
QY 737 PSTIYQE--LQSIKRXH 751
DB 487 FAEIHDDCLQILAETH 502

Search completed: March 18, 2005, 16:00:10
Job time : 50 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 15:41:45 ; Search time 170 Seconds
(without alignments)
1710.846 Million cell updates/sec

Title: US-10-660-763-2

Perfect score: 3879
Sequence: 1 MGFSEELCSPOGHGVLCQWQ.....QRPSTIYQELQSRKRR 752

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3879	100.0	752	5 ABB07354	Abb07354 Human pro
2	3879	100.0	752	6 ABB99046	Abb99046 Human kin
3	3879	100.0	752	8 ADM28581	Adm28581 Human pro
4	3834	98.8	822	5 ABB07355	Abb07355 Human hom
5	3834	98.8	822	6 ABB99047	Abb99047 Human kin
6	3834	98.8	822	8 ADM28583	Adm28583 Human V-F
7	3783.5	97.5	2202	4 AAU32655	Aau32655 Novel hum
8	3493	90.0	764	8 ADR39770	Adr39770 Human kin
9	3473	89.5	820	8 ADH43090	Adh43090 CRM re1a
10	2405	62.0	2017	4 AAU32654	Aau32654 Novel hum
11	2088	53.8	472	8 ADK71829	Adk71829 Human kin
12	1853	47.8	822	7 ADL71055	Adl71055 Type II c
13	1826.5	47.1	823	7 ADL71053	Adl71053 Type II c
14	1364.5	35.2	262	2 AAU33963	Aau33963 Human pro
15	1350.5	34.8	262	2 AAU33964	Aau33964 Cat prote
16	1206.5	31.1	804	4 ABB64045	Abb64045 Drosophi1
17	1130	28.9	211	1 AAP70055	Aap70055 Fes/fps p
18	1116.5	28.8	262	2 AAU33965	Aau33965 Chicken p
19	1006	25.9	256	1 AAP60010	Aap60010 Sequence
20	884	22.8	592	4 ABB66018	Abb66018 Drosophi1
21	861	17.0	258	5 ADL17265	Adl17265 Polypepti
22	661	17.0	258	5 ADL17322	Adl17322 Polypepti
23	657	16.9	256	5 ABB78820	Abb78820 Tyrosine
24	657	16.9	257	5 ABB78796	Abb78796 Tyrosine
25	655	16.9	2039	8 ADR06335	Adr06335 Gsk-3 bet

ALIGNMENTS

RESULT 1
ID ABB07354 standard; protein; 752 AA.
XX
AC ABB07354;
XX
DT 09-APR-2002 (first entry)
XX
DE Human proto-oncogene tyrosine kinase.
XX
KW Proto-oncogene tyrosine kinase; poTK; tumour; cytostatic; anti-leukemic;
KW gene therapy; protein therapy; vaccine; enzymatic-inhibition; human;
KW anti-kinase.
XX
OS Homo sapiens.
XX
PN US6340584-B1.
XX
PD 22-JUN-2002.
XX
PF 27-MAR-2001; 2001US-00817180.
XX
PR 27-MAR-2001; 2001US-00817180.
XX
PS (PEKE) PE CORP NY.
XX
PT Gan W, Ye J, Di Francesco V, Beasley EM;
XX
PI WPI; 2002-138497/18.
XX
PS N-PSDB; ABA94500, ABA94501.
XX
PT Nucleic acids encoding a proto-oncogene tyrosine kinase, useful for the
XX prevention, diagnosis and treatment of e.g. leukemia and lung tumors.
XX
PS Claim 1; Fig 2A; 49pp; English.
XX
CC The invention provides isolated nucleic acid sequences encoding a proto-
CC oncogene tyrosine kinase (poTK). The poTK polynucleotides and protein may
CC be used in the prevention, diagnosis and treatment of diseases associated
CC with inappropriate poTK expression, such as lung and kidney tumours,
CC leukemia and stomach adenocarcinoma. poTK may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of poTK by expressing
CC inactive proteins or to supplement the patient's own production of poTK.
CC The encoded poTK may be used as an antigen in the production of poTK
CC antibodies against poTK and in assays to identify modulators of poTK
CC expression and activity. The anti-poTK antibodies and antagonists may be
CC used to down regulate expression and activity and as diagnostic agents


```

QY 241 VVAIHREMAAAARIOPEAEYOGFLROYGSADVPVPCVTFDESILLEGEPLEBGELOLNE 300
DB 241 VVAIHREMAAAARIOPEAEYOGFLROYGSADVPVPCVTFDESILLEGEPLEBGELOLNE 300
QY 301 LTVESVOHTLTVSTDELAVALATEMVFRRQEMVTQLOQELNBEENTHPRRVOVLGKROYL 360
DB 301 LTVESVOHTLTVSTDELAVALATEMVFRRQEMVTQLOQELNBEENTHPRRVOVLGKROYL 360
QY 361 QEALQGLQVALCSQAKLOAQOQELQTKLEHLGEPPEPVLLLODDHSTSSSEQEREGGR 420
DB 361 QEALQGLQVALCSQAKLOAQOQELQTKLEHLGEPPEPVLLLODDHSTSSSEQEREGGR 420
QY 421 TPTLEILKSHIGSIFRPKPSNLYRLEGEGPSPILIDHLSTQPLTKKSGVHLRAVP 480
DB 421 TPTLEILKSHIGSIFRPKPSNLYRLEGEGPSPILIDHLSTQPLTKKSGVHLRAVP 480
QY 481 KDKWVLMHEDLVLEQOIGRNGFGEVFSGRLRADNTLVAVKSCRETLPPDLKAKFLOEARI 540
DB 481 KDKWVLMHEDLVLEQOIGRNGFGEVFSGRLRADNTLVAVKSCRETLPPDLKAKFLOEARI 540
QY 541 LKQYSHPNIVRLIGVCTQKOPITYIVMELVQGGDFLTPLETEGARLVKTLQWVGDAAG 600
DB 541 LKQYSHPNIVRLIGVCTQKOPITYIVMELVQGGDFLTPLETEGARLVKTLQWVGDAAG 600
QY 601 MEYLESKCCIHRLAARNCLVTEKNVLTSDFGMSREEDGVYAAGSLRQVPVKMTABE 660
DB 601 MEYLESKCCIHRLAARNCLVTEKNVLTSDFGMSREEDGVYAAGSLRQVPVKMTABE 660
QY 661 ALNYGRYSSESDVWSFGILLMETFSLGASPPYPLNSQOTREFEVEKGRLLPCBELCPDAVF 720
DB 661 ALNYGRYSSESDVWSFGILLMETFSLGASPPYPLNSQOTREFEVEKGRLLPCBELCPDAVF 720
QY 721 RLMEQCMAYEPGQRPFSFTIYQELQSIKRRHR 752
DB 721 RLMEQCMAYEPGQRPFSFTIYQELQSIKRRHR 752

RESULT 3
ADM28581
ID ADM28581 standard; protein; 752 AA.
XX
AC ADM28581;
XX
DT 17-JUN-2004 (first entry)
XX
DE Human proto-oncogene tyrosine kinase.
XX
KW gene therapy; human; proto-oncogene; tyrosine kinase; cancer; enzyme.
XX
OS Homo sapiens.
XX
PN US2004063130-A1.
XX
PD 01-APR-2004.
XX
PF 12-SEP-2003; 2003US-00660763.
XX
PR 27-MAR-2001; 2001US-00817180.
PR 06-DEC-2001; 2001US-00003295.
XX
PA (APPL-) APPLERA CORP.
XX
PI Gan W, Ye J, Di Francesco V, Beasley EM;
XX
DR N-PSDB; ADM28580, ADM28582.
XX
PT New human kinase peptides, useful for preparing a composition for
XX treating a disease or condition mediated by human kinases.
XX PS Claim 1; SEQ ID NO 2; 53pp; English.
XX CC The invention relates to a new isolated human proto-oncogene tyrosine

```

```

CC kinase peptide or its allelic variant, orthologue or fragment. The
CC peptide is useful for preparing a composition for treating a disease or
CC condition mediated by a human kinase protein e.g. cancer. The present
CC sequence represents the amino acid sequence of the human proto-oncogene
CC tyrosine kinase.
XX
SQ Sequence 752 AA;

Query Match 100.0%; Score 3879; DB 8; Length 752;
Best Local Similarity 100.0%; Pred. No. 6e-280;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSSSELSPQOHGVLQMOQABRLRLLEGMRKMAQVRSVDEYAGLHHMSIQDSGGOS 60
DB 1 MEFSSSELSPQOHGVLQMOQABRLRLLEGMRKMAQVRSVDEYAGLHHMSIQDSGGOS 60
QY 61 RAISPSPISQSWAETSQTGELSRLLRQHADLNSGPLSKSLIRERQRLKTYSEOM 120
DB 61 RAISPSPISQSWAETSQTGELSRLLRQHADLNSGPLSKSLIRERQRLKTYSEOM 120
QY 121 QOLQOELTTHSODIEKLSQYRALARDSAOKKRYQASKDKDKADKQYVRSIMKLF 180
DB 121 QOLQOELTTHSODIEKLSQYRALARDSAOKKRYQASKDKDKADKQYVRSIMKLF 180
QY 181 AHHNRVVLGVRAQOLHHQHQLLLPGLRSLQDLHEEACILKETLOEYLSISLVODE 240
DB 181 AHHNRVVLGVRAQOLHHQHQLLLPGLRSLQDLHEEACILKETLOEYLSISLVODE 240
QY 241 VVAIHREMAAAARIOPEAEYOGFLROYGSADVPVPCVTFDESILLEGEPLEBGELOLNE 300
DB 241 VVAIHREMAAAARIOPEAEYOGFLROYGSADVPVPCVTFDESILLEGEPLEBGELOLNE 300
QY 301 LTVESVOHTLTVSTDELAVALATEMVFRRQEMVTQLOQELNBEENTHPRRVOVLGKROYL 360
DB 301 LTVESVOHTLTVSTDELAVALATEMVFRRQEMVTQLOQELNBEENTHPRRVOVLGKROYL 360
QY 361 QEALQGLQVALCSQAKLOAQOQELQTKLEHLGEPPEPVLLLODDHSTSSSEQEREGGR 420
DB 361 QEALQGLQVALCSQAKLOAQOQELQTKLEHLGEPPEPVLLLODDHSTSSSEQEREGGR 420
QY 421 TPTLEILKSHIGSIFRPKPSNLYRLEGEGPSPILIDHLSTQPLTKKSGVHLRAVP 480
DB 421 TPTLEILKSHIGSIFRPKPSNLYRLEGEGPSPILIDHLSTQPLTKKSGVHLRAVP 480
QY 481 KDKWVLMHEDLVLEQOIGRNGFGEVFSGRLRADNTLVAVKSCRETLPPDLKAKFLOEARI 540
DB 481 KDKWVLMHEDLVLEQOIGRNGFGEVFSGRLRADNTLVAVKSCRETLPPDLKAKFLOEARI 540
QY 541 LKQYSHPNIVRLIGVCTQKOPITYIVMELVQGGDFLTPLETEGARLVKTLQWVGDAAG 600
DB 541 LKQYSHPNIVRLIGVCTQKOPITYIVMELVQGGDFLTPLETEGARLVKTLQWVGDAAG 600
QY 601 MEYLESKCCIHRLAARNCLVTEKNVLTSDFGMSREEDGVYAAGSLRQVPVKMTABE 660
DB 601 MEYLESKCCIHRLAARNCLVTEKNVLTSDFGMSREEDGVYAAGSLRQVPVKMTABE 660
QY 661 ALNYGRYSSESDVWSFGILLMETFSLGASPPYPLNSQOTREFEVEKGRLLPCBELCPDAVF 720
DB 661 ALNYGRYSSESDVWSFGILLMETFSLGASPPYPLNSQOTREFEVEKGRLLPCBELCPDAVF 720
QY 721 RLMEQCMAYEPGQRPFSFTIYQELQSIKRRHR 752
DB 721 RLMEQCMAYEPGQRPFSFTIYQELQSIKRRHR 752

RESULT 4
ABB07355
ID ABB07355 standard; protein; 822 AA.
XX
AC ABB07355;
XX
DT 09-APR-2002 (first entry)
XX

```

QY		36	QELALGSLGVALLGSQAKLQAQOELLQTKEHLKHPGPSPVLLIQDRHSTSSSEDEREGGR	420
Db		361	QELAGSLGVALLGSQAKLQAQOELLQTKEHLKHPGPSPVLLIQDRHSTSSSEDEREGGR	420
QY		421	TPTLEILKSHISGIIFRPKS-----	440
Db		421	TPTLEILKSHISGIIFRPKFSLLPPLQLIPEVOKRLHEQLWTHGAIPRAEVAELVHSGRF	480
QY		441	-----NLVLRGGGPSIFLLIDHLISTQQPLTKK	470
Db		481	LVRSEGSGKOEYVLSQLMDGLPRHFITIGSLDMLVRLGGGPSIFLLIDHLISTQQPLTKK	540
QY		471	SGVLTARAVPKDKWVLNHHEDLVLGEOIRGNFGVEVFSGRLRADNTLVAAVKSCRETLPDDL	530
Db		541	SGVLTARAVPKDKWVLNHHEDLVLGEOIRGNFGVEVFSGRLADNTLVAAVKSCRETLPDDL	600
QY		531	KAKFLQEARILKQYSHPNIVRLIGVCTOKOPRIYYIMELVQGSDTLTFRTEGARLARVKTL	590
Db		601	KAKFLQEARILKQYSHPNIVRLIGVCTOKOPRIYYIMELVQGSDTLTFRTEGARLARVKTL	660
QY		591	LQWVGDAAGAAMEYLEBSKCCIRHDLAARNCLVTENKVLKISDFGMSREBADVVAASGCLR	650
Db		661	LQWVGDAAGAAMEYLESKCCIRHDLAARNCLVTENKVLKISDFGMSREBADVVAASGCLR	720
QY		651	QVEPVKTATBEALNYGYSSESDDWSFGILLWETSLSGASPYNLSNOOTREFVEKGRLP	710
Db		721	QVEPVKTATBEALNYGYSSESDDWSFGILLWETSLSGASPYNLSNOOTREFVEKGRLP	780
QY		711	CPELCPDAAVFRIMEOCMAYPEGPCSFSTIYOELOSTRKRRR	752
Db		781	CPELCPDAAVFRIMEOCMAYPEGPCSFSTIYOELOSTRKRRR	822
RESULT 5				
ABB99047	ID	ABB99047	standard; protein; 822 AA.	
XX	AC	ABB99047;		
XX	DT	24-JAN-2003	(first entry)	
XX	XX		Human kinase related amino acid sequence.	
XX	HM		Human; kinase; proto-oncogene tyrosine kinase; antiarteriosclerotic;	
XX	KM		cytostatic; antinflammatory; antiproliferative; gene therapy;	
XX	KM		protein kinase; drug screening assay; tissue typing; chromosome 15;	
XX	KM		pharmacogenomic analysis; inflammation; cancer; leukaemia; lung tumours;	
XX	KX		kidney tumour; stomach adenocarcinoma; arteriosclerosis; psoriasis.	
OS	XX	Homo sapiens.		
XX	PN	WO200277191-A2.		
XX	PD	03-OCT-2002.		
XX	PF	27-MAR-2002; 2002MO-US009325.		
XX	PR	27-MAR-2001; 2001US-00817180.		
XX	PR	06-DEC-2001; 2001US-00003295.		
XX	PA	(PEKE) PE CORP NY.		
XX	FI	Gan W, Ye J, Di Francesco V, Beasley EM;		
XX	DR	WPI; 2003-029927/02.		
PT	XX	New human kinase protein, useful for treating or diagnosing disorders		
PT	XX	associated with an absence of, inappropriate, or unwanted expression of		
PT	XX	the protein, e.g. inflammation or cancer, in drug screening assays and		
PT	XX	pharmacogenomics.		
XX	XX	Disclosure; Page 74-75; 75pp; English.		

CC The invention relates to a newly isolated peptide sequence of a human
 CC kinase that is related to the proto-oncogene tyrosine kinase subfamily.
 CC The activity of the kinase of the invention may be described as,
 CC cytoactive, antiarteriosclerotic, antiinflammatory and antiproliferative.
 CC Peptides of the invention are useful in assays to determine the
 CC biological activity of the protein, in drug screening assays, tissue
 CC typing and pharmacogenomic analysis. They are also useful in treating or
 CC diagnosing disorders characterized by an absence of, inappropriate, or
 CC unwanted expression of the protein, such as inflammation, cancer (e.g.
 CC leukemia, lung tumours, kidney tumours or stomach adenocarcinoma),
 CC arteriosclerosis, and psoriasis. Nucleic acid molecules of the invention
 CC are useful as probes, primers and chemical intermediates in biological
 CC assays. The peptide and nucleic acid sequences are useful as models for
 CC the development of human therapeutic targets, aid in the identification
 CC of therapeutic proteins and serve as targets for the development of human
 CC therapeutic agents that modulate kinase protein activity in cells and
 CC tissues that express the protein. The protein of the invention may also
 CC be useful in gene therapy. The gene encoding the protein of the invention
 CC has been localised to human chromosome 15. The current sequence
 CC represents an amino acid sequence relative to the human kinase sequence
 CC of the invention. NOTE: This sequence is not further mentioned in the
 CC specification
 CC XX
 CC XX

SO Sequence 822 AA;

Query Match 98.8%; Score 3834; DB 6; Length 822;

Best Local Similarity 91.5%; Pred. No. 1.5e-276;

Matches 752; Conservative 0; Mismatches 0; Indels 70; Gaps 1;

QY 1 MGFSSELCSPOGHGVLCQOEAEIRLLBGRKMAQVKSDBREYAGLLHMSLQDSGGOS 60
 DB 1 MGFSSELCSPOGHGVLCQOEAEIRLLBGRKMAQVKSDBREYAGLLHMSLQDSGGOS 60
 QY 61 RAISPSPISSQSWAETTSQTGSLRLRQAEEDLNSGPKSLILRERQQLKRTYSEOW 120
 DB 61 RAISPSPISSQSWAETTSQTGSLRLRQAEEDLNSGPKSLILRERQQLKRTYSEOW 120
 QY 121 QOLQOELTKTHSGDIEKLKSQYRALARDSQAQARKYQEAASKDORPKADKYVRSIMKLF 180
 DB 121 QOLQOELTKTHSGDIEKLKSQYRALARDSQAQARKYQEAASKDORPKADKYVRSIMKLF 180
 QY 121 QOLQOELTKTHSGDIEKLKSQYRALARDSQAQARKYQEAASKDORPKADKYVRSIMKLF 180
 DB 121 QOLQOELTKTHSGDIEKLKSQYRALARDSQAQARKYQEAASKDORPKADKYVRSIMKLF 180
 QY 181 AHNHRVYLVGRAQLHNOHNOHLLPGLARSLODLHEEMACIKETLOELETSSLYODE 240
 DB 181 AHNHRVYLVGRAQLHNOHNOHLLPGLARSLODLHEEMACIKETLOELETSSLYODE 240
 QY 241 VVAIHREMAAAARIQPEAEYQGLRQYGSAPVPCVTFDESILBEGEPLEPGELOANE 300
 DB 241 VVAIHREMAAAARIQPEAEYQGLRQYGSAPVPCVTFDESILBEGEPLEPGELOANE 300
 QY 301 LTVESVQHTLTSTYDELAVATEVFRROEMWTOLQOELRNEBENTHPRERVOQLGRQVL 360
 DB 301 LTVESVQHTLTSTYDELAVATEVFRROEMWTOLQOELRNEBENTHPRERVOQLGRQVL 360
 QY 361 QEALQGLQVALGSOAKLQAOQELLOTLEHLGEGEPVYLLDDDHSTSSSOBERGGR 420
 DB 361 QEALQGLQVALGSOAKLQAOQELLOTLEHLGEGEPVYLLDDDHSTSSSOBERGGR 420
 QY 421 TPTLEILKSHISGIFRPKFS----- 440
 DB 421 TPTLEILKSHISGIFRPKFS----- 440
 QY 441 -----NLYRLEGEGFPSPILLIDHLSTFOPLTKK 470
 DB 481 LVRESQKQKQYVLLMDGLPRHFIIOQLNLVRLBEGFPSPILLIDHLSTFOPLTKK 540
 QY 471 SGVVLRAVPKDKWLVNHEDLVYGEOTGRNGFVSGRLRADNTLVAVVSCETLPDIL 560
 DB 541 SGVVLRAVPKDKWLVNHEDLVYGEOTGRNGFVSGRLRADNTLVAVVSCETLPDIL 600
 QY 531 KAKFLEARILKQYSHPNIVRLIGVCTQKQPIYIVMELVQSGDFLFLRTGARLRKTKL 590
 DB 601 KAKFLEARILKQYSHPNIVRLIGVCTQKQPIYIVMELVQSGDFLFLRTGARLRKTKL 660

QY 591 LQWVGDAAGMEYTESKCIHRDLARNCLVTEKNVLTISDEGMSREBAGVYAAAGGLR 650
 DB 661 LQWVGDAAGMEYTESKCIHRDLARNCLVTEKNVLTISDEGMSREBAGVYAAAGGLR 720
 QY 651 QVPVKMTAPEALNAGRYSESDVWSPGILLMETFSLGASPPYNI,SNQOTREPVKGGRLP 710
 DB 721 QVPVKMTAPEALNAGRYSESDVWSPGILLMETFSLGASPPYNI,SNQOTREPVKGGRLP 780
 QY 711 CPFLCPDAVFRIMEQCMAYEPGQRPSPFTTYOELQIRKRRH 752
 DB 781 CPFLCPDAVFRIMEQCMAYEPGQRPSPFTTYOELQIRKRRH 822

RESULT 6

ID ADM28583 standard; protein; 822 AA.

AC ADM28583;

DT 17-JUN-2004 (first entry)

DE Human V-FBS/FPS protein.

KW Gene therapy; human; proto-oncogene; tyrosine kinase; cancer; enzyme;

KM V-FBS/FPS.

OS Homo sapiens.

PN US2004063130-A1.

PD 01-APR-2004.

PF 12-SEP-2003; 2003US-00660763.

PR 27-MAR-2001; 2001US-00817180.

PR 06-DEC-2001; 2001US-00003295.

PA (APPL-) APPLERA CORP.

PI Gan W, Ye J, Di Francesco V, Beasley EM;

DR WPI; 2004-282461/26.

PT New human kinase peptides, useful for preparing a composition for

PS treating a disease or condition mediated by human kinases.

PS Disclosure; SEQ ID NO 4; 53pp; English.

CC The invention relates to a new isolated human proto-oncogene tyrosine

CC kinase peptide or its allelic variant, orthologue or fragment. The

CC peptide is useful for preparing a composition for treating a disease or

CC condition mediated by a human kinase protein e.g. cancer. The present

CC sequence represents the amino acid sequence of the human V-FBS/FPS

CC protein.

CC XX

CC XX

SO Sequence 822 AA;

Query Match 98.8%; Score 3834; DB 8; Length 822;

Best Local Similarity 91.5%; Pred. No. 1.5e-276;

Matches 752; Conservative 0; Mismatches 0; Indels 70; Gaps 1;

QY 1 MGFSSELCSPOGHGVLCQOEAEIRLLBGRKMAQVKSDBREYAGLLHMSLQDSGGOS 60
 DB 1 MGFSSELCSPOGHGVLCQOEAEIRLLBGRKMAQVKSDBREYAGLLHMSLQDSGGOS 60
 QY 61 RAISPSPISSQSWAETTSQTGSLRLRQAEEDLNSGPKSLILRERQQLKRTYSEOW 120
 DB 61 RAISPSPISSQSWAETTSQTGSLRLRQAEEDLNSGPKSLILRERQQLKRTYSEOW 120
 QY 121 QOLQOELTKTHSGDIEKLKSQYRALARDSQAQARKYQEAASKDORPKADKYVRSIMKLF 180
 DB 121 QOLQOELTKTHSGDIEKLKSQYRALARDSQAQARKYQEAASKDORPKADKYVRSIMKLF 180

```

QY 181 AHNHRYVLGVRAAQLHHQHLLPGLRLSLQDLHEENACTIKELIQEYLEISLVQDE 240
DB 181 AHNHRYVLGVRAAQLHHQHLLPGLRLSLQDLHEENACTIKELIQEYLEISLVQDE 240
QY 241 VVAIHREMAAAARIQPEAEYQGFRLQYGSAPVPCVTFDESLLEEGEPLEBGEQLQNE 300
DB 241 VVAIHREMAAAARIQPEAEYQGFRLQYGSAPVPCVTFDESLLEEGEPLEBGEQLQNE 300
QY 301 LVVESVQHTLTSVTDELAVALATEMVFRRQEMVNTQLOQLNEENHTPRERVOLGKROYL 360
DB 301 LVVESVQHTLTSVTDELAVALATEMVFRRQEMVNTQLOQLNEENHTPRERVOLGKROYL 360
QY 361 QEALQGLQVALCSQAKLQAOQELQTKLEHLGPGEPPVLLQDDHSTSSSQEREGGR 420
DB 361 QEALQGLQVALCSQAKLQAOQELQTKLEHLGPGEPPVLLQDDHSTSSSQEREGGR 420
QY 421 TPTLEILKSHISGIFRPKFS----- 440
DB 421 TPTLEILKSHISGIFRPKFSLPPLQLIPEVQKPLHEQLMYGALPRAEVAELLVHSGDF 480
QY 441 -----NLVRLSEGGPSPILIDHLSTQOPLTKK 470
DB 481 LVRESQKQBYVLVSLMDGLPRHFIITQSLDNLVRLSEGGPSPILIDHLSTQOPLTKK 540
QY 471 SGVVLHRAVPKDKMVLNHEDLVLEGOIGRGNFGEVSGRLRADNTLVAVKSCRETLPPDL 530
DB 541 SGVVLHRAVPKDKMVLNHEDLVLEGOIGRGNFGEVSGRLRADNTLVAVKSCRETLPPDL 600
QY 531 KAKFLQBARILKQYSHPNVRLIGVCTOKOPYITVWELVQGGDFLPLRTGEGARLVKTL 590
DB 601 KAKFLQBARILKQYSHPNVRLIGVCTOKOPYITVWELVQGGDFLPLRTGEGARLVKTL 660
QY 591 LQWVGDAAGMEYLESKCCIHRLDAAARNCLVTEKNVLTISDFGMSREBAGVVAASGGR 650
DB 661 LQWVGDAAGMEYLESKCCIHRLDAAARNCLVTEKNVLTISDFGMSREBAGVVAASGGR 720
QY 721 QVAVKMTAPBALNYGRYSSESDVMSFGILLMETFSLGASPYPLSNQOTREFVEKGRLLP 710
DB 721 QVAVKMTAPBALNYGRYSSESDVMSFGILLMETFSLGASPYPLSNQOTREFVEKGRLLP 780
QY 711 CPFLCPDAVFRMLQOCMAVEPGQRPSPSTTYQSLQIRKRR 752
DB 781 CPFLCPDAVFRMLQOCMAVEPGQRPSPSTTYQSLQIRKRR 822

RESULT 7
AAU32655 standard; protein; 2202 AA.
ID AAU32655;
AC AAU32655;
XX
XX 18-DEC-2001 (first entry)
DT
XX
XX Novel human secreted protein #3146.
DE
XX
XX Human; vaccination; gene therapy; nutritional supplement;
KM stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
OS Homo sapiens.
PN WO200179449-A2.
XX
XX 25-OCT-2001.
PD
XX
XX 16-APR-2001; 2001MO-US008656.
PF
XX
XX 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-0070160.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Drmanac RT;
PI

```

```

XX WPI; 2001-611725/70.
DR
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
XX Claim 20; Page 645-646; 765pp; English.
PS
XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU32655-2001 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 2202 AA;
Query Match 97.5%; Score 3783.5; DB 4; Length 2202;
Best Local Similarity 84.3%; Pred. No. 3,4e-272;
Matches 751; Conservative 0; Mismatches 1; Indels 139; Gaps 2;
QY 1 MGFSSELCSPOGHGVLQOQWAEFLLEGRKMAQRYVSDREYAGLHHMSLQDSGQGS 60
DB 1 MGFSSELCSPOGHGVLQOQWAEFLLEGRKMAQRYVSDREYAGLHHMSLQDSGQGS 60
QY 61 RAISPDSPISQMAEITQTEGLSRLRQHAEDLNSGPIKSLILIREQOQRKTYSEGM 120
DB 61 RAISPDSPISQMAEITQTEGLSRLRQHAEDLNSGPIKSLILIREQOQRKTYSEGM 120
QY 121 QOLOELTHTHODIEKLSQYRALARDAQAKRYQESKQKDRKADKXVRSIMKLF 180
DB 121 QOLOELTHTHODIEKLSQYRALARDAQAKRYQESKQKDRKADKXVRSIMKLF 180
QY 121 QOLOELTHTHODIEKLSQYRALARDAQAKRYQESKQKDRKADKXVRSIMKLF 180
DB 121 QOLOELTHTHODIEKLSQYRALARDAQAKRYQESKQKDRKADKXVRSIMKLF 180
QY 181 AHNHRYVLGVRAAQLHHQHLLPGLRLSLQDLHEENACTIKELIQEYLEISLVQDE 240
DB 181 AHNHRYVLGVRAAQLHHQHLLPGLRLSLQDLHEENACTIKELIQEYLEISLVQDE 240
QY 241 VVAIHREMAAAARIQPEAEYQGFRLQYGSAPVPCVTFDESLLEEGEPLEBGEQLQNE 300
DB 241 VVAIHREMAAAARIQPEAEYQGFRLQYGSAPVPCVTFDESLLEEGEPLEBGEQLQNE 300
QY 301 LVVESVQHTLTSVTDELAVALATEMVFRRQEMVNTQLOQLNEENHTPRERVOLGKROYL 360
DB 301 LVVESVQHTLTSVTDELAVALATEMVFRRQEMVNTQLOQLNEENHTPRERVOLGKROYL 360
QY 361 QEALQGLQVALCSQAKLQAOQELQTKLEHLGPGEPPVLLQDDHSTSSSQEREGGR 420
DB 361 QEALQGLQVALCSQAKLQAOQELQTKLEHLGPGEPPVLLQDDHSTSSSQEREGGR 420
QY 421 TPTLEILKSHISGIFRPKFS----- 440
DB 421 TPTLEILKSHISGIFRPKFSLPPLQLIPEVQKPLHEQLMYGALPRAEVAELLVHSGDF 480
QY 441 -----NLVRLSEGGPSPILIDHLSTQOPLTKK 470
DB 481 LVRESQKQBYVLVSLMDGLPRHFIITQSLDNLVRLSEGGPSPILIDHLSTQOPLTKK 540
QY 453 IPLLIDHLSTQOPLTKKSGVVLHRAVPKDKMVLNHEDLVLEGOIGR 499
DB 541 IPLLIDHLSTQOPLTKKSGVVLHRAVPKDKMVLNHEDLVLEGOIGR 600
QY 500 -----GNFGEVFGRLRADNTLVAVKS 521

```


DB 601 RGNTPGHPGVSRMGRKRRLRMWEGRSPRPQGNFSGVSGRLRADNTLVAVNS 660
 QY 522 CRETLPDDKAKFLQARILKQYSHPNIVRLIGVCTOKQPIYIMELVQGGPFLRLTE 581
 DB 661 CRETLPDDKAKFLQARILKQYSHPNIVRLIGVCTOKQPIYIMELVQGGPFLRLTE 720
 QY 582 GARLPVKTLLQWVGDAAGMEYLSKCCIHRLDAANCILTEKNVTKIDPFGSREAAQ 641
 DB 721 GARLPVKTLLQWVGDAAGMEYLSKCCIHRLDAANCILTEKNVTKIDPFGSREAAQ 780
 QY 642 VVAASGGLRQVPVKMTAPALNNGRYSSESDDVMSFGILLMETFSLGASPYPLNSQOTRE 701
 DB 781 VVAASGGLRQVPVKMTAPALNNGRYSSESDDVMSFGILLMETFSLGASPYPLNSQOTRE 840
 QY 702 FVEKGRLLPCPELCPDAVFLMEQCAVYBPGRPSFTIYQELQIRKRR 752
 DB 841 FVEKGRLLPCPELCPDAVFLMEQCAVYBPGRPSFTIYQELQIRKRR 891
 RESULT 8
 ADR39770 standard; protein; 764 AA.
 ADR39770;
 18-NOV-2004 (first entry)
 Human kinase and phosphatase KRP-43 protein SEQ ID NO:43.
 human; kinase and phosphatase protein; KRP; enzyme; cytosolic;
 antileukemic; anticonvulsant; neuroprotective;
 cerebroprotective; anti-HIV; antiallergic; antiinflammatory;
 chymotrypsin; gene therapy; cell proliferative disorder; cancer;
 atherosclerosis; neurological disorder; epilepsy; Huntington's disease;
 stroke; immune disorder; inflammatory disorder; AIDS; allergy;
 developmental disorder; Hypothyroidism; Cushing's syndrome; infection;
 KRP-43.
 KM
 OS Homo sapiens.
 PN MO2004074453-AZ.
 XX
 PD 02-SEP-2004.
 XX
 PF 20-FEB-2004; 2004MO-US05092.
 XX
 PR 20-FEB-2003; 2003US-0449059P.
 PR 19-MAR-2003; 2003US-0456932P.
 PR 28-MAR-2003; 2003US-0458844P.
 PR 09-APR-2003; 2003US-0461678P.
 PR 17-APR-2003; 2003US-0463937P.
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Rankumar J, Margulis JP, Swarnakar A, Chawla NK, Tran UK,
 PI Becha SO, Lee SY, Hafalia AJA, Richardson TW, Khare R, Jiang X,
 PI Jackson AA, Yang J, Gorvad AE;
 XX
 DR MPI: 2004-635568/61.
 DR N-PSDB; ADR39816.
 XX
 PT New human kinases and phosphatases (KRP) for diagnosing, treating and
 PT preventing diseases or conditions associated with aberrant KRP expression
 PT e.g. cancer, acquired immunodeficiency syndrome, epilepsy, or infections.
 XX
 PS Claim 1; SEQ ID NO 43; 299pp; English.
 XX
 CC The present sequence represents the human kinase and phosphatase protein
 CC (KRP), designated KRP-43. The human KRP sequences from the present
 CC invention have cytosolic, antileukemic, anticonvulsant,
 CC neuroprotective, cerebroprotective, anti-HIV, antiallergic,
 CC antiinflammatory and chymotrypsin activities, and can be used in gene

CC therapy. The human KRP proteins and polynucleotides can be used in
 CC diagnosing, treating and preventing diseases or conditions associated
 CC with the decreased expression or overexpression of KRP, such as cell
 CC proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
 CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
 CC allergies) and developmental (e.g. Hypothyroidism, Cushing's syndrome)
 CC disorders, or infections. They can also be used in assessing the effects
 CC of exogenous compounds on the expression of nucleic acid and amino acid
 CC sequences of KRP. The KRP or its fragments are useful in screening
 CC compounds for effectiveness as agonist or antagonist of the polypeptides,
 CC or in altering the expression of the target polynucleotide and compounds
 CC that specifically bind to or modulate the activity of the polypeptide.
 XX
 SQ Sequence 764 AA;
 Query Match 90.0%; Score 3493; DB 8; Length 764;
 Best Local Similarity 84.1%; Pred. No. 3.5e-251;
 Matches 691; Conservative 2; Mismatches 1; Indels 128; Gaps 2;
 QY 1 MGFSSSLCPQGHGVLQWQWAEALRLLEGMRKMAQRYKSDREYAGLLHMSLDQSGGOS 60
 DB 1 MGFSSSLCPQGHGVLQWQWAEALRLLEGMRKMAQRYKSDREYAGLLHMSLDQSGGOS 60
 QY 61 RAISPSPISQSWAITSQTEGLSRLLRQHARDLNSGPKSLILREHQRLKTYSEQM 120
 DB 61 RAISPSPISQ----- 71
 QY 121 QQLQQLTKTMSODIEKLKQRYALARDSAQAARKYQASQKDRKADKRYRSLMKLF 180
 DB 72 -----THSDIEKLKQRYALARDSAQAARKYQASQKDRKADKRYRSLMKLF 122
 QY 181 AHNHRYVLGVRAAQLHGHQHLPLGLRSIQDHEBACTLKELQYELISIVQDE 240
 DB 122 AHNHRYVLGVRAAQLHGHQHLPLGLRSIQDHEBACTLKELQYELISIVQDE 182
 QY 241 VVAIHREMAAARLOPEAEYQFLROYGSAADVPVCPVDSLESGPLSPGELQUNE 300
 DB 183 VVAIHREMAAARLOPEAEYQFLROYGSAADVPVCPVDSLESGPLSPGELQUNE 242
 QY 301 LTVESVQHTLTSVTDELAVALTEVFRROEMVTOLOQELNEBENTHPRERVOLGKQVYL 360
 DB 243 LTVESVQHTLTSVTDELAVALTEVFRROEMVTOLOQELNEBENTHPRERVOLGKQVYL 302
 QY 361 QEALQGLQVALCSQAQLAQOQELQTKLEHLPGEPPLVLLQODRHSTSSSQBEREGR 420
 DB 303 QEALQGLQVALCSQAQLAQOQELQTKLEHLPGEPPLVLLQODRHSTSSSQBEREGR 362
 QY 421 TPTLILKSHISGIRPKFS----- 440
 DB 363 TPTLILKSHISGIRPKFSLEPPQLIPEVQKPLHEQLMYHGAIPRAVAELVHSQDF 422
 QY 441 -----NLYRLEGSGPSPILLIDHLLSTQOPLTKK 470
 DB 423 LVRESQKQEVYLVLMQGLPRHFIIGSLNDLYRLEGSGFPSPILLIDHPLSTQOPLTKK 482
 QY 471 SGVTLHRAVPKQKRWLNHEDVLVGEQIGRNGEVPFGSLRADNTLVAVKSGRETLPPL 530
 DB 483 SGVTLHRAVPKQKRWLNHEDVLVGEQIGRNGEVPFGSLRADNTLVAVKSGRETLPPL 542
 QY 531 KAKPLOBARILKQYSPNIVRLIGVCTOKQPIYIMELVQGGPFLRLTEGARLVKTL 590
 DB 543 KAKPLOBARILKQYSPNIVRLIGVCTOKQPIYIMELVQGGPFLRLTEGARLVKTL 602
 QY 591 LQWVGDAAGMEYLSKCCIHRLDAARNCILTEKNVTKIDPFGSREAAQGVVAASGGLR 650
 DB 603 LQWVGDAAGMEYLSKCCIHRLDAARNCILTEKNVTKIDPFGSREAAQGVVAASGGLR 662
 QY 651 QVPVKMTAPALNNGRYSSESDDVMSFGILLMETFSLGASPYPLNSQOTREFEKGRLLP 710
 DB 663 QVPVKMTAPALNNGRYSSESDDVMSFGILLMETFSLGASPYPLNSQOTREFEKGRLLP 722
 QY 711 CPELCPDAVFLMEQCAVYBPGRPSFTIYQELQIRKRR 752

XX Claim 20, Page 645; 765pp; English.

CC The invention relates to novel human secreted polypeptides: The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation, to regulate haematopoiesis, and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention

XX Sequence 2017 AA;

Query Match 62.0%; Score 2405; DB 4; Length 2017;
Best Local Similarity 73.1%; Pred. No. 1,6e-169; Indels 180; Gaps 3;
Matches 488; Conservative 0; Mismatches 0;

QY 224 KEILQYELISSIVQDEVAIHEMAAAARIQPEAEYQGLRQYSSADVPVCTFDES 283
DB 1 KEILQYELISSIVQDEVAIHEMAAAARIQPEAEYQGLRQYSSADVPVCTFDES 60
QY 284 LEEGEPLEPGELOLNELVESVQHTLTSTDELAIVATEVPRKQEMVTOLOQLNEEBE 343
DB 61 LEEGEPLEPGELOLNELVESVQHTLTSTDELAIVATEVPRKQEMVTOLOQLNEEBE 86
QY 344 NTHPRRVOVLGKRYQALQGLQVALCSQAKLQAOELLQTKLEHLGEGEPVTLIQ 403
DB 87 -----VOLLGRQVLOERLQGLQVALCSQAKLQAOELLQTKLEHLGEGEPVTLIQ 139
QY 404 DDRHSTSSSEQEREGRTPTLEILKSHIGIFRPKFS----- 440
DB 140 DDRHSTSSSEQEREGRTPTLEILKSHIGIFRPKFSLEPPQLIPEVQKPLHEQLWYHG 199
QY 441 ----- 440
DB 200 AIPRAEVAELVHSGDFLVRESQKQEVYLSVLMDELPRHFIQSLDGSRLPMEADPG 259
QY 441 -----NLVRLLEGSGFPSPILLIDHLSTOQPLTKSGVVLHRAVPRKDWLNHEDLVGE 495
DB 260 SPALQNLVRLLEGSGFPSPILLIDHLSTOQPLTKSGVVLHRAVPRKDWLNHEDLVGE 319
QY 496 QIGR-----GNFGS 504
DB 320 QIGRVPQRGSNSGRVAVRGPNGTAPHPGVGSRKGRKRRBELRWEGRSPRPQGNFGE 379
QY 505 VFSGRIRADNTLVAVKSCRETLPPDLKAFLOEARILKQYSHNRYRLIGVCTQOKPIYI 564
DB 380 VFSGRIRADNTLVAVKSCRETLPPDLKAFLOEARILKQYSHNRYRLIGVCTQOKPIYI 439
QY 565 VNELVGGDFLPTLRTEGARLRYKTLLOWGDAAGMEYTESCCJHRDLAANCLVTEK 624
DB 440 VNELVGGDFLPTLRTEGARLRYKTLLOWGDAAGMEYTESCCJHRDLAANCLVTEK 499
QY 625 NVLKISDFGMSREADGVYAAASGLRQVPKWTAPALNNGRYSSESDVWSFGILLMEF 684
DB 500 NVLKISDFGMSREADGVYAAASGLRQVPKWTAPALNNGRYSSESDVWSFGILLMEF 559
QY 685 SUGASYPNLSNQGTETPEVEKGRLLPCPELCPDAVRLMEQCAAYEPGQRPSTTYOEL 744
DB 560 SUGASYPNLSNQGTETPEVEKGRLLPCPELCPDAVRLMEQCAAYEPGQRPSTTYOEL 619
QY 745 QSIRKRRH 752
DB 745 QSIRKRRH 752

DB 620 QSIRKRRH 627

RESULT 11
ADK71829
ID ADK71829 standard; protein; 472 AA.
XX
AC ADK71829;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human kinase and phosphatase KPP-6 protein.
XX
KW human; kinase; phosphatase; KPP; cardiovascular; antiarteriosclerotic;
KW hypotensive; vasotropic; antiinflammatory; antianginal; anti-HIV;
KW antiallergic; antiasthmatic; immunosuppressive; antithyroid;
KW dermatological; antidiabetic; nephrotropic; angiotensin; gastroenteric;
KW neuroprotective; osteopathic; antiarthritic; uropathic; ophthalmological;
KW antirheumatic; antiparkinsonian; nootropic; anticonvulsant; hepatotropic;
KW antipneumatic; haemostatic; cytostatic; antilipemic; antiparasitic;
KW antihelminthic; antibacterial; virucide; protozoacide; fungicide;
KW cardiovascular disease; immune system; neurological; growth; development;
KW cell proliferation; viral; bacterial; fungal; parasitic; protozoan;
KW helminthic infection; transgenic; gene therapy; enzyme;
KW single nucleotide polymorphism; SNP.
XX
OS Homo sapiens.
PN WO2004018641-A2.
XX
PD 04-MAR-2004.
XX
PE 25-AUG-2003; 2003WO-US026635.
XX
PR 26-AUG-2002; 2002US-0406172P.
PR 25-SEP-2002; 2002US-0413910P.
PR 27-SEP-2002; 2002US-0414296P.
PR 11-OCT-2002; 2002US-0417821P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Baughn MR, Richardson TW, Marguis JP, Swarnakar A, Tang YT;
PI Becha SD, Emerling BM, Jin P, Wilson AD, Yue H, Gietzen KJ;
PI Chang H, Yang YG, Lee SY, Khare R, Elliott VS, Hafalla AJA;
PI Chawla NK, Ramkumar J, Gururajan R, Tribouley CM, Chien D, Tran UK;
PI Mudge J;
XX
DR WPI: 2004-226830/21.
DR N-PSDB; ADK71888.
XX
PT New human kinases and phosphatases, useful for diagnosing, treating or
PT preventing atherosclerosis, hypertension, AIDS, allergy, multiple
PT sclerosis, osteoarthritis, Alzheimer's disease, Crohn's disease, cancer
PT or hepatitis.
XX
PS Claim 1; SEQ ID NO 6; 347bp; English.
XX
XX The invention relates to a novel isolated polypeptide which is a human
CC kinase and phosphatase (KPP). The polypeptide of the invention
CC demonstrates cardiovascular, antiarteriosclerotic, hypotensive,
CC vasotropic, antiinflammatory, antianginal, anti-HIV, antiallergic,
CC antipneumatic, immunosuppressive, antithyroid, dermatological,
CC antidiabetic, nephrotropic, angiotensin, gastroenteric, neuroprotective,
CC osteopathic, antirheumatic, uropathic, ophthalmological, antiparasitic,
CC antiparkinsonian, nootropic, anticonvulsant, hepatotropic, antihelminthic,
CC haemostatic, cytostatic, antilipemic, antiparasitic, antihelminthic,
CC antibacterial, virucide, protozoacide and fungicide activities. The
CC kinase and phosphatase (KPP) polynucleotides, polypeptides, agonists and
CC antagonists may be useful for diagnosing, treating or preventing
CC disorders such as cardiovascular diseases, immune system disorders,
CC neurological disorders, disorders affecting growth and development, cell
CC proliferative disorders and viral, bacterial, fungal, parasitic, cell
CC protozoan or helminthic infections. Furthermore, the molecules of the

CC Invention may be useful for creating transgenic animals to model human
CC disease and during gene therapy. The current sequence is that of a human
CC KRP protein of the invention.

XX Sequence 472 AA;

Query Match 53.8%; Score 2088; DB 8; Length 472;

Best Local Similarity 99.8%; Pred. No. 8, 8e-147;

Matches 411; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGFSSELCPQGHGVLQOQOEALRLLEGKRMMAQRVKSDEYAGLLHMSLQDSGGQS 60

DB 1 MGFSSELCPQGHGVLQOQOEALRLLEGKRMMAQRVKSDEYAGLLHMSLQDSGGQS 60

QY 61 RAISPSPIQSQAETTSQTEGLSRLLRQHAEDLNSGPIKSLRLRERQRLKTYSEGM 120

DB 61 RAISPSPIQSQAETTSQTEGLSRLLRQHAEDLNSGPIKSLRLRERQRLKTYSEGM 120

QY 121 QOLQOELTTHSQDIEKLKSOYRALARDSQAQKRYQASKODRDXKDKYVRSIMKLF 180

DB 121 QOLQOELTTHSQDIEKLKSOYRALARDSQAQKRYQASKODRDXKDKYVRSIMKLF 180

QY 181 AHNRYVYLVGRAAQLHGHQHQLLPGILRLSLQDLHEMACILKEILOYLEISLVDDE 240

DB 181 AHNRYVYLVGRAAQLHGHQHQLLPGILRLSLQDLHEMACILKEILOYLEISLVDDE 240

QY 241 VVAIHREMAAARIQPEAEYQGLROYGAPDVPFCVTFDESLLGEGBLEGELOLNE 300

DB 241 VVAIHREMAAARIQPEAEYQGLROYGAPDVPFCVTFDESLLGEGBLEGELOLNE 300

QY 301 LTVESVQHTLTSVTDLAVALATEMVFRRQEWVTOLOQELRNEBENTHPRERVOQLGKROYL 360

DB 301 LTVESVQHTLTSVTDLAVALATEMVFRRQEWVTOLOQELRNEBENTHPRERVOQLGKROYL 360

QY 361 QEALQGLQVALCSQAQLQAOQELLOTKLHLGEPPEPVLLLODDRHSSTSS 412

DB 361 QEALQGLQVALCSQAQLQAOQELLOTKLHLGEPPEPVLLLODDRHSSTSS 412

RESULT 12

ADL71055 standard; protein; 822 AA.

XX ADL71055;

DT 20-MAY-2004 (first entry)

XX Type II collagen expression promoting protein, seq id 48.

XX Osteopathic; antiinflammatory; antiarthritis; antiarthritic;

KW gene therapy; type II collagen; expression; cartilage disease;

KW osteoarthritis; cartilage defect; rheumatoid arthritis; human.

OS Homo sapiens.

PN WO2003087375-A1.

PD 23-OCT-2003.

PF 16-APR-2003; 2003WO-0P04802.

PR 16-APR-2002; 2002JP-0013908.

PR 19-APR-2002; 2002US-0373594P.

PA (ASAH) ASAH KASEI KK.

PI Matsuda A., Honda G., Muramatsu S;

PT New purified protein that promotes type II collagen expression, useful for preventing and treating a cartilage disease, e.g. osteoarthritis,

PT cartilage defect, or rheumatoid arthritis.

XX Claim 1, SEQ ID NO 48; 271pp; English.

XX The invention relates to a purified protein (I) that promotes type II

CC collagen expression. Also disclosed is an isolated polynucleotide

CC encoding (I), a recombinant vector comprising the polynucleotide and a

CC gene therapeutic agent comprising the recombinant vector as an active

CC ingredient. The protein, genes, agents and methods are useful for

CC preventing and treating a cartilage disease, e.g. osteoarthritis,

CC cartilage defect, or rheumatoid arthritis. The current sequence

CC represents a human protein that promotes type II collagen expression.

XX Sequence 822 AA;

Query Match 47.8%; Score 1853; DB 7; Length 822;

Best Local Similarity 45.5%; Pred. No. 6, 1e-129; Indels 84; Gaps 9;

DB 376; Conservative 150; Mismatches 217;

QY 1 MGFSSELCPQGHGVLQOQOEALRLLEGKRMMAQRVKSDEYAGLLHMSLQDSGGQS 60

DB 1 MGFSSELCPQGHGVLQOQOEALRLLEGKRMMAQRVKSDEYAGLLHMSLQDSGGQS 60

QY 61 RAISPSPIQSQAETTSQTEGLSRLLRQHAEDLNSGPIKSLRLRERQRLKTYSEGM 120

DB 61 RAISPSPIQSQAETTSQTEGLSRLLRQHAEDLNSGPIKSLRLRERQRLKTYSEGM 120

QY 121 QOLQOELTTHSQDIEKLKSOYRALARDSQAQKRYQASKODRDXKDKYVRSIMKLF 179

DB 121 QOLQOELTTHSQDIEKLKSOYRALARDSQAQKRYQASKODRDXKDKYVRSIMKLF 179

QY 181 AHNRYVYLVGRAAQLHGHQHQLLPGILRLSLQDLHEMACILKEILOYLEISLVDDE 239

DB 181 AHNRYVYLVGRAAQLHGHQHQLLPGILRLSLQDLHEMACILKEILOYLEISLVDDE 239

QY 240 EVVAIHREMAAARIQPEAEYQGLROYGAPDVPFCVTFDESLLGEGBLEGELOLNE 299

DB 240 EVVAIHREMAAARIQPEAEYQGLROYGAPDVPFCVTFDESLLGEGBLEGELOLNE 299

QY 300 ELTVESVQHTLTSVTDLAVALATEMVFRRQEWVTOLOQELRNEBENTHPRERVOQLGKROYL 358

DB 300 ELTVESVQHTLTSVTDLAVALATEMVFRRQEWVTOLOQELRNEBENTHPRERVOQLGKROYL 358

QY 359 VLOEALQGLQVALCSQAQLQAOQELLOTKLHLGEPPEPVLLLODDRHSSTSS 418

DB 359 VLOEALQGLQVALCSQAQLQAOQELLOTKLHLGEPPEPVLLLODDRHSSTSS 418

QY 417 -RLSKFESIRHSIAGIIRSPKSAVSSALSDMISIEKPLAQDWYHGAIPRIEAOELLK 475

DB 417 -RLSKFESIRHSIAGIIRSPKSAVSSALSDMISIEKPLAQDWYHGAIPRIEAOELLK 475

QY 439 -----FSNLYRLEGGEPPSPILLIDHLSNQ 464

DB 439 -----FSNLYRLEGGEPPSPILLIDHLSNQ 464

QY 476 KQGDPLVRESHGKPEBYLVSYSDGQRHFIQYVDNMFRFEGTGSNIPOLLIDHLYTK 535

DB 476 KQGDPLVRESHGKPEBYLVSYSDGQRHFIQYVDNMFRFEGTGSNIPOLLIDHLYTK 535

QY 465 QPLTKKSGVVLRAVAPKD-KWVLANHEDLVAGEIGRGNGEYFSGRLRADNTLVAVKSCR 523

DB 465 QPLTKKSGVVLRAVAPKD-KWVLANHEDLVAGEIGRGNGEYFSGRLRADNTLVAVKSCR 523

QY 536 QVITKKSGLVNLPIPKOKKWTLSHEDVILGELLKGNNGEYVYKGLK-DKTSVAVKTK 594

DB 536 QVITKKSGLVNLPIPKOKKWTLSHEDVILGELLKGNNGEYVYKGLK-DKTSVAVKTK 594

```

Db      774 EGGYRMSAPQHCPEDISKIMKCMWDYKBPENRPFSELOKELTIIRK 820

RESULT 13
ADL71053
ID      ADL71053 standard; protein; 823 AA.
XX
XX      ADL71053;
XX
XX      20-MAY-2004 (first entry)
XX
XX      Type II collagen expression promoting protein, seq id 46.
XX
XX      Osteopathic; antiinflammatory; antirheumatic; antiarthritic;
XX      gene therapy; type II collagen; expression; cartilage disease;
XX      osteoarthritis; cartilage defect; rheumatoid arthritis; human.
XX
XX      Homo sapiens.
XX
XX      WO2003087375-A1.
XX
XX      23-OCT-2003.
XX
XX      16-APR-2003; 2003WO-JP004802.
XX
XX      16-APR-2002; 2002JP-00113908.
XX      19-APR-2002; 2002US-0373594P.
XX
XX      (ASAH ) ASAMI KASEI KK.
XX
XX      Matsuda A, Honda G, Muramatsu S;
XX      WPI; 2003-845331/78.
XX      DR      N-PSDB; ADL71052.
XX
XX      New purified protein that promotes type II collagen expression, useful
XX      for preventing and treating a cartilage disease, e.g. osteoarthritis,
XX      cartilage defect, or rheumatoid arthritis.
XX
XX      Claim 1; SEQ ID NO 46; 271pp; English.
XX
XX      The invention relates to a purified protein (I) that promotes type II
XX      collagen expression. Also disclosed is an isolated polynucleotide
XX      encoding (I), a recombinant vector comprising the polynucleotide and a
XX      gene therapeutic agent comprising the recombinant vector as an active
XX      ingredient. The proteins, genes, agents and methods are useful for
XX      preventing and treating a cartilage disease, e.g. osteoarthritis,
XX      cartilage defect, or rheumatoid arthritis. The current sequence
XX      represents a human protein that promotes type II collagen expression.
XX
XX      Sequence 823 AA;

Query Match      47.1%; Score 1826.5; DB 7; Length 823;
Best Local Similarity 45.5%; Pred. No. 5.8e-127;
Matches 376; Conservative 147; Mismatches 219; Indels 85; Gaps 10;

QY      1 MGFSSSELCPQGHGVLQOMQAEELRLLEGKMKMAQVXSDREYAGLLHMSLQDSGGQS 60
DB      1 MGFSGSLDKNSQ-EAVL-KLQDMELRLLETVKKFMALRIKSDKEYAYTLQNLQNVDKEST 58
QY      61 RAISPPSPISQSWAEITTSQTEGSLRLRQHAEDLNSGPKSLRLREKQQLKRTYSEOW 120
DB      59 VQVNVYNSVSKSLMLIQOTEQSLRIKMTAEDLNSGPHRLTMMIKDQVVKKSYVGIN 118
QY      121 QOLQOELTYTHSDIEKLKSOYRALARDAQAQKTKQEA-SKDQDPAKADKTVRSLMTKL 179
DB      119 QQIEBAMIKVTETELKLSSTYQRLIKEMASAKETKEALAKGETEKAKERTDKATMKL 178
QY      180 FAHHNRVYLVGAQAQLHQNHLQLLPGLLRSIQDLAHEMACTLKELQEYLEISSLYOD 239
DB      179 HMLHNGYVALAKQAQHQSQYVDTTLPLLLDSVQKQOEEMIKALKGIFPDYSQITSLVTR 238
QY      240 EYVAIHRBMAAAAAARIQPRAEYQGFURQGSAPDVPCTTFDESLLEBGEPLRGLQLN 239

```

```

Db      239 EIVVNHKEIQMSVEQIDPESTETNNFLDYHRTTAAKQGELEFDTSLIEENENQANEIMWN 298
QY      300 ELTVESVQHTLTSVTDDELAVATEWVFRQEWNTOLQOELRNEENTHPR-ERVOLLGKQ 358
Db      299 NLTADSLQVMTKLABELTQOTQOMLHKEAAVLIEKRIEESPEFCEKSDIYLLIGQK 358
QY      359 VLOEALQGLQVNLCSQAKLQAOQELQTLLEHLEGPEPPVULLDDRRSTSSSEQEREG 418
Db      359 ALBELKQSVQOLRCTEAKCAQKALLEQKQENDEGPEPPVNVYBEDARSVTSMEKE-- 416
QY      419 GRPTLEIKSHISGIFR-PK----- 438
Db      417 -RLSKFESIRSHIAGIIRKPKSVLSSSTQVCVTSVGERPLAEHDWYHGAIRIEAQELL 475
QY      439 -----FSNLYRLEGSGFSPILPLIDHLLST 463
Db      476 KQGGDFLVNESHGKPEEYVLVSYSDQRRHFLIQFDNLYRFEQTFSNIPOLIDHFMNT 535
QY      464 QOPLTKSGVYLHRAVPKD-KVYLNHEDLVLEQIGRGKFGVPSGRRLADNTLVAVKSC 522
Db      536 KOVITKSGVVLNPIPKDKKWLNHEDVSLGELLGKGFGEVYKGTIK-DKTPVALIKTC 594
QY      523 RETLPDDLKAKFLQEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGDPLTFPRTG 582
Db      595 KEDLPQELKIKFLQEKILKQYDHNIVKLIQVCTQROQVYIIMELVPGDPLTFIRKX 654
QY      583 ARLRVETLLQVNDAAAGMEYLESKCIRHDLAARNCVTEKNVLIKISDFGMSREBADGV 642
Db      655 DELKDKQLVRFSLDVAAGMLVLESKNCHRDLAARNCVGENNTLTKISDFGMSRQDGGV 714
QY      643 YAAAGGLQVYPVYKTAPEALNAGRYSSBSVDVMSFGILMETSLGASPIPNLSNOQTREF 702
Db      715 YSSS-GLKQIPIKWAPALNAGRYSSBSVDVMSFGILMETPSLGVCPYPMGTMOQARQ 773
QY      703 VERGGRGLPCPELCPDAVFRIMECWAYPEGQRPSPFTIVQEOSIRK 749
Db      774 VERGTRMSAPQHCPESEVFTIMKCMWDYKBPENRPFSELOKELTIIRK 820

RESULT 14
AAY43963
ID      AAY43963 standard; protein; 262 AA.
XX
XX      AAY43963;
XX
XX      21-DEC-1999 (first entry)
XX
XX      Human protein kinase #18.
XX
XX      Prediction; secondary structure; alignment; evolutionary conservation;
XX      KW      homology; periodicity; co-variation analysis; antigenic site;
XX      site directed mutagenesis; interaction.
XX
XX      Homo sapiens.
XX
XX      US5958784-A.
XX
XX      28-SEP-1999.
XX
XX      25-MAR-1992; 92US-00857224.
XX
XX      25-MAR-1992; 92US-00857224.
XX
XX      (BENNER) BENNER S A.
XX
XX      Benner SA;
XX
XX      WPI; 1999-570766/48.
XX
XX      Predicting the folded structure of proteins.
XX
XX      Disclosure; Col 271-274; 113pp; English.

```

XX Sequences AAY43902-Y44015 represent proteins used in a novel method of
CC predicting the folded structure of proteins, by aligning sequences of
CC homologous proteins and using patterns of evolutionarily conserved and
CC varied sequences to assign positions. Positions in the alignment are
CC assigned to the surface or inside of the folded structure, active sites,
CC and pairing segments. Secondary structural units are assigned by
CC identifying periodicity in the assignments, and assembled into globular
CC form using distance constraints imposed by disulfide bridges, active site
CC assignments and co-variation analysis. The predicted secondary structures
CC are useful for identifying antigenic sites on a protein molecule, as
CC guides for site directed mutagenesis studies, and for understanding the
CC interaction of a protein with other molecules

XX Sequence 262 AA;

Query Match 35.2%; Score 1364.5; DB 2; Length 262;
Best Local Similarity 99.6%; Pred. No. 3.2e-93;
Matches 262; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

DB 489 EDLVIGEQTIGRNFGEVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKOYSHN 548
1 EDLVIGEQTIGRNFGEVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKOYSHN 60
QY 549 IYRLIGVCTQOKPIYIYVMEIVOGGDFLFLRTGRLRVKTLLOWVGDAAAGMEYLESKC 608
61 IYRLIGVCTQOKPIYIYVMEIVOGGDFLFLRTGRLRVKTLLOWVGDAAAGMEYLESKC 120
DB 609 CIHRDLAARNCLVTEKNVLKISDFGMSREADGVYAASGLRQVPYKWTAPALNNGRYS 668
121 CIHRDLAARNCLVTEKNVLKISDFGMSREADGVYAAS-GLRQVPYKWTAPALNNGRYS 179
QY 669 SESDWSFGILLMETFSLGASPYPNLSNOOTREFVEKGRLLCPCLCPDAVFLMEQCWA 728
DB 180 SESDWSFGILLMETFSLGASPYPNLSNOOTREFVEKGRLLCPCLCPDAVFLMEQCWA 239
QY 729 YEPGQRPSPFTIYQELQSIKRRH 751
DB 240 YEPGQRPSPFTIYQELQSIKRRH 262

RESULT 15

AAY43964
ID AAY43964 standard; protein; 262 AA.

XX AAY43964;

XX 17-OCT-2003 (revised)
DT 21-DEC-1999 (first entry)

XX Cat protein kinase #1.

XX Prediction; secondary structure; alignment; evolutionary conservation;
KW homology; periodicity; co-variation analysis; antigenic site;
KW site directed mutagenesis; interaction.

XX Fells catus.

XX US5958784-A.

XX 28-SEP-1999.

XX 25-MAR-1992; 92US-00857224.

XX 25-MAR-1992; 92US-00857224.

XX (BENN/) BENNER S A.

XX Benner SA;

XX WPI, 1999-570766/48.

XX Predicting the folded structure of proteins.

XX PS Disclosure; Col 275-276; 113pp; English.

XX Sequences AAY43902-Y44015 represent proteins used in a novel method of
CC predicting the folded structure of proteins, by aligning sequences of
CC homologous proteins and using patterns of evolutionarily conserved and
CC varied sequences to assign positions. Positions in the alignment are
CC assigned to the surface or inside of the folded structure, active sites,
CC and pairing segments. Secondary structural units are assigned by
CC identifying periodicity in the assignments, and assembled into globular
CC form using distance constraints imposed by disulfide bridges, active site
CC assignments and co-variation analysis. The predicted secondary structures
CC are useful for identifying antigenic sites on a protein molecule, as
CC guides for site directed mutagenesis studies, and for understanding the
CC interaction of a protein with other molecules. (updated on 17-OCT-2003 to
CC standardise OS field)

XX Sequence 262 AA;

Query Match 34.8%; Score 1350.5; DB 2; Length 262;
Best Local Similarity 97.7%; Pred. No. 3.5e-92;
Matches 257; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 489 EDLVIGEQTIGRNFGEVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKOYSHN 548
1 EDLVIGEQTIGRNFGEVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKOYSHN 60
DB 549 IYRLIGVCTQOKPIYIYVMEIVOGGDFLFLRTGRLRVKTLLOWVGDAAAGMEYLESKC 608
61 IYRLIGVCTQOKPIYIYVMEIVOGGDFLFLRTGRLRVKTLLOWVGDAAAGMEYLESKC 120
DB 609 CIHRDLAARNCLVTEKNVLKISDFGMSREADGVYAASGLRQVPYKWTAPALNNGRYS 668
121 CIHRDLAARNCLVTEKNVLKISDFGMSREADGVYAAS-GLRQVPYKWTAPALNNGRYS 179
QY 669 SESDWSFGILLMETFSLGASPYPNLSNOOTREFVEKGRLLCPCLCPDAVFLMEQCWA 728
DB 180 SESDWSFGILLMETFSLGASPYPNLSNOOTREFVEKGRLLCPCLCPDAVFLMEQCWA 239
QY 729 YEPGQRPSPFTIYQELQSIKRRH 751
DB 240 YEPGQRPSPFTIYQELQSIKRRH 262

Search completed: March 18, 2005, 15:59:19
Job time : 176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 10:14:07 ; Search time 1174.22 Seconds
(without alignments)
11437.798 Million cell updates/sec

Title: US-10-660-763-1_COPY_72_2327

Perfect score: 2256
Sequence: 1 atggagcttccttcctgagct.....gacccgaagcgagctcg 2256

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 554486 seqs, 2976611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- Published Applications NA:*
- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
 - 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
 - 10: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
 - 11: /cgn2_6/prodata/2/pubpna/US09C_NEW_PUB.seq:*
 - 12: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
 - 13: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
 - 14: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
 - 15: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
 - 16: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
 - 17: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
 - 18: /cgn2_6/prodata/2/pubpna/US10F_PUBCOMB.seq:*
 - 19: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
 - 20: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
 - 21: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
 - 22: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2256	100.0	2674	13 US-10-003-295-1	Sequence 1, Appl1
2	2256	100.0	2674	17 US-10-660-763-1	Sequence 1, Appl1
3	2023.4	89.7	2889	16 US-10-240-965-256	Sequence 256, App
4	326.8	14.5	449	10 US-09-918-995-1503	Sequence 1503, Ap
5	296.2	13.1	361	9 US-09-948-802-5	Sequence 5, Appl1
6	296.2	13.1	361	9 US-10-121-925-5	Sequence 5, Appl1
7	271.6	12.0	1779	17 US-10-280-576-19	Sequence 19, Appl1
8	227.2	10.1	15297	13 US-10-003-295-3	Sequence 3, Appl1
9	227.2	10.1	15297	17 US-10-660-763-3	Sequence 3, Appl1
10	201.2	8.9	3875	15 US-10-101-510-525	Sequence 525, App
11	198	8.8	2955	18 US-10-384-339C-1	Sequence 1, Appl1

12	198	8.8	3370	9 US-09-967-768A-144	Sequence 144, App
13	198	8.8	3370	16 US-10-354-356-101	Sequence 101, App
14	198	8.8	3370	16 US-10-210-120-19	Sequence 19, Appl
15	198	8.8	3370	18 US-10-776-827-82	Sequence 82, Appl
16	198	8.8	3370	18 US-10-473-974-219	Sequence 219, App
17	185.4	8.2	3042	18 US-10-384-339C-2	Sequence 2, Appl
18	185.4	8.2	3921	10 US-09-921-406C-29	Sequence 29, Appl
19	185.4	8.2	3921	15 US-10-007-926A-222	Sequence 222, App
20	185.4	8.2	3921	16 US-10-269-909-19	Sequence 19, App
21	185.4	8.2	3921	17 US-10-366-288-1	Sequence 1, Appl1
22	185.4	8.2	3921	17 US-10-172-118-1036	Sequence 1036, Ap
23	185.4	8.2	3921	17 US-10-342-887-1036	Sequence 1036, Ap
24	185.4	8.2	3921	17 US-10-287-226-319	Sequence 319, App
25	185.4	8.2	3921	18 US-10-648-593-1	Sequence 1, Appl1
26	185.4	8.2	3921	18 US-10-473-974-223	Sequence 223, App
27	185.4	8.2	3921	19 US-10-616-403-5	Sequence 5, Appl1
28	185.4	8.2	3935	19 US-10-897-711-1	Sequence 1, Appl1
29	176.2	7.8	1518	17 US-10-280-576-24	Sequence 24, Appl
30	174.6	7.7	816	18 US-10-377-268-3	Sequence 3, Appl1
31	174.6	7.7	1050	18 US-10-377-268-4	Sequence 4, Appl1
32	174.6	7.7	3030	16 US-10-325-430-8	Sequence 8, Appl1
33	174.6	7.7	3416	8 US-08-987-689A-1	Sequence 1, Appl1
34	174.6	7.7	3416	15 US-10-292-524-1	Sequence 1, Appl1
35	174.6	7.7	3416	17 US-10-464-805-2	Sequence 2, Appl1
36	174.6	7.7	3416	17 US-10-305-720-1483	Sequence 1483, Ap
37	174.6	7.7	3544	14 US-10-161-803-51	Sequence 51, Appl
38	174.6	7.7	4089	17 US-10-291-808-13	Sequence 13, Appl
39	174.6	7.7	4089	18 US-10-620-052A-7	Sequence 7, Appl1
40	174.6	7.7	4151	14 US-10-161-803-52	Sequence 52, Appl
41	174.6	7.7	4151	16 US-10-325-430-7	Sequence 7, Appl1
42	170.4	7.6	1611	18 US-10-476-962-3	Sequence 3, Appl1
43	170.4	7.6	1611	18 US-10-887-588-1	Sequence 1, Appl1
44	170.4	7.6	1611	18 US-10-887-588-3	Sequence 3, Appl1
45	170.4	7.6	1611	19 US-10-497-641-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-10-003-295-1
; Sequence 1, Application US/10003295
; Publication NO. US20020168741A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THREEOF
; FILE REFERENCE: CLO01183DIY
; CURRENT APPLICATION NUMBER: US/10/003,295
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO. 1
; LENGTH: 2674
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-003-295-1

Query Match	100.0%	Score 2256	DB 13	Length 2674
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2256	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	ATGGGCTTCCTCTGAGCTGTGACAGCCGAGGCGGAGTCTCTGACGAAATGCAG	60	
DB	72	ATGGGCTTCCTCTGAGCTGTGACAGCCGAGGCGGAGTCTCTGACGAAATGCAG	131	
QY	61	GAGGCGGAGTCTCTGAGCTGTGACAGGCGGAGGCGGAGTCTCTGAGGCGGAGTCT	120	
DB	132	GAGGCGGAGTCTCTGAGCTGTGACAGGCGGAGGCGGAGTCTCTGAGGCGGAGTCT	191	
QY	121	GACAGGAGATGACAGGAGTCTTCCATGTCCTGACGAGGAGGCGGAGGCGGAGG	180	

Db 192 GACAGGAGATGATGACGAGTCTTTCACCAATGTCCTTGACGAGACAGTGGGGCCAGAGC 251
Qy 181 CGGGCCATCAGCCCTGACAGCCCATTCAGTCACTCTGGGCTGAGATCACACCAACT 240
Db 252 CGGGCCATCAGCCCTGACAGCCCATTCAGTCACTCTGGGCTGAGATCACACCAACT 311
Qy 241 GAGGGCTGAGCCGCTGCTGCGGACGACGACAGAGATCTGAATCAGGGCCCTGAGC 300
Db 312 GAGGGCTGAGCCGCTGCTGCGGACGACGACAGAGATCTGAATCAGGGCCCTGAGC 371
Qy 301 AAGCTGAGCTGCTCATCCGGGAAACGGACAGCTTTCGAAAGCTTACAGCCAGCATG 360
Db 372 AAGCTGAGCTGCTCATCCGGGAAACGGACAGCTTTCGAAAGCTTACAGCCAGCATG 431
Qy 361 CAGCACTGAGAGAGAGCTCACCAAGACCCACAGCCAGCATTTAGAGAGCTGAAAGC 420
Db 432 CAGCACTGAGAGAGAGCTCACCAAGACCCACAGCCAGCATTTAGAGAGCTGAAAGC 491
Qy 421 CAGTACCGAGCTCTGGACCGGACAGTGCACCAAGCCAAAGCCAAAGTACAGAGCCAGC 480
Db 492 CAGTACCGAGCTCTGGACCGGACAGTGCACCAAGCCAAAGCCAAAGTACAGAGCCAGC 551
Qy 481 AAGACAAGAGACCGTGAACAAGCCAAAGAGATGTCGACAGCTGTGAAAGCTCTTT 540
Db 552 AAGACAAGAGACCGTGAACAAGCCAAAGAGATGTCGACAGCTGTGAAAGCTCTTT 611
Qy 541 GCTCACCAACAACGGCTATGTCGTGGGCTGGGGCTGGAGCTTACCAACCAAGCAAC 600
Db 612 GCTCACCAACAACGGCTATGTCGTGGGCTGGGGCTGGAGCTTACCAACCAAGCAAC 671
Qy 601 CACAGAGCTCTGCTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 672 CACAGAGCTCTGCTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 731
Qy 661 TGCATCTGAAGAGATCTGTCAGAAATACCTGAGATGACAGCTGTGTCAGATGAG 720
Db 732 TGCATCTGAAGAGATCTGTCAGAAATACCTGAGATGACAGCTGTGTCAGATGAG 791
Qy 721 GTGGTGGCCATTCACCGGAGATGAGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 792 GTGGTGGCCATTCACCGGAGATGAGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTG 851
Qy 781 TACCAAGGCTTCTGCGACAGTATGAGTCCGACCTGACCTGACCTGCTGCTGCTGCT 840
Db 852 TACCAAGGCTTCTGCGACAGTATGAGTCCGACCTGACCTGACCTGCTGCTGCTGCT 911
Qy 841 GATGAGTCACTGCTTGAAGAGAGTGAACGCTGAGAGCTTGGAGAGCTTCAAGTGAAG 900
Db 912 GATGAGTCACTGCTTGAAGAGAGTGAACGCTGAGAGCTTGGAGAGCTTCAAGTGAAG 971
Qy 901 CTGACTGTGAGAGCGTGCAGACACGCTGACCTGACCTGACCTGACCTGACCTGACCT 960
Db 972 CTGACTGTGAGAGCGTGCAGACACGCTGACCTGACCTGACCTGACCTGACCTGACCT 1031
Qy 961 ACCGAGATGATGTTCAAGCGGACAGAGATGATGATGATGATGATGATGATGATGAT 1020
Db 1032 ACCGAGATGATGTTCAAGCGGACAGAGATGATGATGATGATGATGATGATGATGAT 1091
Qy 1021 GAAAGAGAGAACCCACCCCGGAGAGCGAGTGCAGTCTGTGGCAAGAGAGAGAGTCTG 1080
Db 1092 GAAAGAGAGAACCCACCCCGGAGAGCGAGTGCAGTCTGTGGCAAGAGAGAGAGTCTG 1151
Qy 1081 CAAGAGACATGCAAGGAGCTGCAAGATGAGCTGTGCAAGCCAGAGCTGCAAGCCAG 1140
Db 1152 CAAGAGACATGCAAGGAGCTGCAAGATGAGCTGTGCAAGCCAGAGCTGCAAGCCAG 1211
Qy 1141 CAGAGTTGCTGCAAGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTG 1200
Db 1212 CAGAGTTGCTGCAAGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTG 1271
Qy 1201 CTCCTGAGAGATGACCGCACTCCACGTCGTCTCTGAGAGAGAGAGAGAGAGAGAG 1260
Db 1272 CTCCTGAGAGATGACCGCACTCCACGTCGTCTCTGAGAGAGAGAGAGAGAGAGAG 1331

Qy 1261 AACCCAGCGTGGAGATCTTAAAGCCACATCTCAGAAATCTTCCGCCCAAGTTCTCG 1320
Db 1332 AACCCAGCGTGGAGATCTTAAAGCCACATCTCAGAAATCTTCCGCCCAAGTTCTCG 1391
Qy 1321 AACCTGTACCGAATGGAAGGAGAGCTTCTTCTGAGATCTTCTGCTCATGACCTTA 1380
Db 1392 AACCTGTACCGAATGGAAGGAGAGCTTCTTCTGAGATCTTCTGCTCATGACCTTA 1451
Qy 1381 CTGAGACCCAGACAGCCCTTCAACCAAGAGATGATGATGATGATGATGATGATGATG 1440
Db 1452 CTGAGACCCAGACAGCCCTTCAACCAAGAGATGATGATGATGATGATGATGATGATG 1511
Qy 1441 AAGGACAAATGGAGTGTGCTGAACCATGAGACCTGATGATGATGATGATGATGATGATG 1500
Db 1512 AAGGACAAATGGAGTGTGCTGAACCATGAGACCTGATGATGATGATGATGATGATGATG 1571
Qy 1501 AACTTTGAGGAAATGTTTCAGCGGACGCTGCGAGCCGACCAACACCTGTGTGGGTGAAG 1560
Db 1572 AACTTTGAGGAAATGTTTCAGCGGACGCTGCGAGCCGACCAACACCTGTGTGGGTGAAG 1631
Qy 1561 TCTTTGCGAGAGACGCTCCCACTTGAACCTCAAGGCTTAAGATTTCTAAGAGAGAGATC 1620
Db 1632 TCTTTGCGAGAGACGCTCCCACTTGAACCTCAAGGCTTAAGATTTCTAAGAGAGAGATC 1691
Qy 1621 CTGAAGCAGTACAGCCACCCCAACATCGTGCCTCAATTTGATGATGATGATGATGATGAT 1680
Db 1692 CTGAAGCAGTACAGCCACCCCAACATCGTGCCTCAATTTGATGATGATGATGATGATGAT 1751
Qy 1681 CCCATCTACATCGTCAATGAGCTTGTGACAGGGGGGAGACTTCTGACCTTCTCCGACG 1740
Db 1752 CCCATCTACATCGTCAATGAGCTTGTGACAGGGGGGAGACTTCTGACCTTCTCCGACG 1811
Qy 1741 GAGGGGGCCCCCTGGGGGTGAAGACTCTGCTGCAAGTGTGGGGGATGACAGCTGCTGCG 1800
Db 1812 GAGGGGGCCCCCTGGGGGTGAAGACTCTGCTGCAAGTGTGGGGGATGACAGCTGCTGCG 1871
Qy 1801 ATGGAATCCTGGAAGAGAAATGCTGATCCACCGGAGACTTGGATCTGGAATCTGCTG 1860
Db 1872 ATGGAATCCTGGAAGAGAAATGCTGATCCACCGGAGACTTGGATCTGGAATCTGCTG 1931
Qy 1861 GTGACAGAGAGAAATGCTGATCCGTAAGATCACTGATGATGATGATGATGATGATGATG 1920
Db 1932 GTGACAGAGAGAAATGCTGATCCGTAAGATCACTGATGATGATGATGATGATGATGATG 1991
Qy 1921 GGGGTATATGACAGCTCAGGGGGCTCAGACAAATGCTGATGATGATGATGATGATGATG 1980
Db 1992 GGGGTATATGACAGCTCAGGGGGCTCAGACAAATGCTGATGATGATGATGATGATGATG 2051
Qy 1981 GCCCTTAATCAAGCGGCTACTCTCCGAAAGGAGAGTGTGATGATGATGATGATGATGATG 2040
Db 2052 GCCCTTAATCAAGCGGCTACTCTCCGAAAGGAGAGTGTGATGATGATGATGATGATGATG 2111
Qy 2041 TGGGAGACCTTCAAGCTGAGGGGCTTCCCTTATCCCACTTCAAGATCAGACACAGG 2100
Db 2112 TGGGAGACCTTCAAGCTGAGGGGCTTCCCTTATCCCACTTCAAGATCAGACACAGG 2171
Qy 2101 GAGTTGTGAGAAAGGGGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Db 2172 GAGTTGTGAGAAAGGGGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2231
Qy 2161 AGGCTCATGAGAGAGTGTGGGCTTATGAGCTGTGGGAGAGGAGGAGGAGGAGGAGGAG 2220
Db 2232 AGGCTCATGAGAGAGTGTGGGCTTATGAGCTGTGGGAGAGGAGGAGGAGGAGGAGGAG 2291
Qy 2221 TACCAAGAGCTGCAAGATCCGAAAGCGGCAATCG 2256
Db 2292 TACCAAGAGCTGCAAGATCCGAAAGCGGCAATCG 2327

RESULT 2
us-10-660-763-1
; Sequence 1, Application US/10660763


```

; Publication No. US20040063130A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C000183DIV1
; CURRENT APPLICATION NUMBER: US/10/660,763
; CURRENT FILING DATE: 2003-09-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2674
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-660-763-1

Query Match      100.0%; Score 2256; DB 17; Length 2674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGGGCTTCTCTCTGAGCTGTGACGCCCCCAGGCGCCACGCGGTCTCTGACGAAATGACG 60
DB      72 ATGGGCTTCTCTCTGAGCTGTGACGCCCCCAGGCGCCACGCGGTCTCTGACGAAATGACG 131
QY      61 GAGGCGGAGTTGCTCTACTGAGGCGATGAAAGTGAATGGCCGCGGTCTCAAGAT 120
DB      132 GAGGCGGAGTTGCTCTACTGAGGCGATGAAAGTGAATGGCCGCGGTCTCAAGAT 191
QY      121 GACAGGAGATGACAGGAGCTTCAACATGTCCTTGAGGACAGTGGGGCCAGAGC 180
DB      192 GACAGGAGATGAGAGAGCTTCAACATGTCCTTGAGGACAGTGGGGCCAGAGC 251
QY      181 CGGCGCATGACCCCTGACAGCCCATCATGTCAGTCTGGGCTGAGATCACAGCCAACT 240
DB      252 CGGCGCATGACCCCTGACAGCCCATCATGTCAGTCTGGGCTGAGATCACAGCCAACT 311
QY      241 GAGGCGCTGAGCGGCTTCTGCGGACGACGACAGAGATTGAATCAGGGCCCTGAGC 300
DB      312 GAGGCGCTGAGCGGCTTCTGCGGACGACGACAGAGATTGAATCAGGGCCCTGAGC 371
QY      301 AAGCTGAGCTGCTCATCGGGAAACGGGACAGCTTGGCAAGACTTACAGGAGAGTGG 360
DB      372 AAGCTGAGCTGCTCATCGGGAAACGGGACAGCTTGGCAAGACTTACAGGAGAGTGG 431
QY      361 CAGCAGCTGACAGAGAGCTTACCAAGACCACAGCCAGGACATTGAGAACTGGAAGC 420
DB      432 CAGCAGCTGACAGAGAGCTTACCAAGACCACAGCCAGGACATTGAGAACTGGAAGC 491
QY      421 CAGTACGAGCTCTGGACAGGAGACGTGCCAAGCCAAAGTACCAAGAGGCCAGC 480
DB      492 CAGTACGAGCTCTGGACAGGAGACGTGCCAAGCCAAAGTACCAAGAGGCCAGC 551
QY      481 AAAGACAAGGACGGTGAACAGGACCAAGATATGTGGAGCCCTGTGAAAGCTTTT 540
DB      552 AAAGACAAGGACGGTGAACAGGACCAAGATATGTGGAGCCCTGTGAAAGCTTTT 611
QY      541 GCTTACCAACAACGCTTATGTGTGGGCGTGGGCTTGGAGCTTACCAACAACAAC 600
DB      612 GCTTACCAACAACGCTTATGTGTGGGCGTGGGCTTGGAGCTTACCAACAACAAC 671
QY      601 CACCAAGCTCTGCTGCGCGGCTGCTGGGTCACTGACGAGACCTGACGAGAGATGGCT 660
DB      672 CACCAAGCTCTGCTGCGCGGCTGCTGGGTCACTGACGAGACCTGACGAGATGGCT 731
QY      661 TGCATCTGAAGAGATCTGACAGAAATACCTGGAATTAGCAAGCCGTGAGAGATGAG 720
DB      732 TGCATCTGAAGAGATCTGACAGAAATACCTGGAATTAGCAAGCCGTGAGAGATGAG 791
QY      721 GTGTGGCCATTACCGGAGATGGCTGACGCTGCGCATCCAGCTGAGGCTGAG 780
DB      792 GTGTGGCCATTACCGGAGATGGCTGACGCTGCGCATCCAGCTGAGGCTGAG 851

```

```

QY      781 TACCAAGCTTCTCTGAGAGATATGGGTCCGACCTGACGTCCACCTGTGTACGCTTC 840
DB      852 TACCAAGCTTCTCTGAGAGATATGGGTCCGACCTGACGTCCACCTGTGTACGCTTC 911
QY      841 GATGAGTCACTGCTTGAAGAGGATGAACGCTGAGACCTGGGAGGCTGACGCTGAACGAG 900
DB      912 GATGAGTCACTGCTTGAAGAGGATGAACGCTGAGACCTGGGAGGCTGACGCTGAACGAG 971
QY      901 CTGACTGTGAGAGCGGTGACGACACGCTGACCTCAAGTACAGATGAGCTGCTGTGGCC 960
DB      972 CTGACTGTGAGAGCGGTGACGACACGCTGACCTCAAGTACAGATGAGCTGCTGTGGCC 1031
QY      961 ACCGAGATGTGTTCAGGCGGCGAGAGATGTTTACGACGTGCAAGAGAGCTCCGGAAT 1020
DB      1032 ACCGAGATGTGTTCAGGCGGCGAGAGATGTTTACGACGTGCAAGAGAGCTCCGGAAT 1091
QY      1021 GAAGAGAGAAACCCACCCCGGAGCGGATGACAGCTGCTGGGCAAGAGGCAAGTCTG 1080
DB      1092 GAAGAGAGAAACCCACCCCGGAGCGGATGACAGCTGCTGGGCAAGAGGCAAGTCTG 1151
QY      1081 CAAGAAACACTGAGGAGGCTGACAGTACGCTGTGACAGCCCAAGCTGACAGGCCAG 1140
DB      1152 CAAGAAACACTGAGGAGGCTGACAGTACGCTGTGACAGCCCAAGCTGACAGGCCAG 1211
QY      1141 CAGGATTTGTGAGAACAAAGCTGAGACCTGTGGGCCCCGGCGAGCCCCGCTGTGCTG 1200
DB      1212 CAGGATTTGTGAGAACAAAGCTGAGACCTGTGGGCCCCGGCGAGCCCCGCTGTGCTG 1271
QY      1201 CTCCTGAGGATGACGCGCACTCCAGTGTCTCTGAGAGAGAGGAGGAGGAGG 1260
DB      1272 CTCCTGAGGATGACGCGCACTCCAGTGTCTCTGAGAGAGAGGAGGAGGAGG 1331
QY      1261 ACAACCAAGCTGAGATCTTTAAGACCAATCTCAGAAATCTTCGCCCCAAGTTCTCG 1320
DB      1332 ACAACCAAGCTGAGATCTTTAAGACCAATCTCAGAAATCTTCGCCCCAAGTTCTCG 1391
QY      1321 AACCTTACCGACTGGAAGGGAAGGCTTCTTACGATTCCTTGTCTCATCCACCTTA 1380
DB      1392 AACCTTACCGACTGGAAGGGAAGGCTTCTTACGATTCCTTGTCTCATCCACCTTA 1451
QY      1381 CTGAGCACCCAGAGCCCTTCAACAAGAGAGTGTGTCTCTGACAGAGGCTGTGCCC 1440
DB      1452 CTGAGCACCCAGAGCCCTTCAACAAGAGAGTGTGTCTCTGACAGAGGCTGTGCCC 1511
QY      1441 AAGACAAGTGGGTGTGAACATGAGACCTGTGTGTGGTGAAGACATTTGAACGAGG 1500
DB      1512 AAGACAAGTGGGTGTGAACATGAGACCTGTGTGTGGTGAAGACATTTGAACGAGG 1571
QY      1501 AACTTTGGGAAAGTGTTCAGCGGACGCTGCGAGCCGACAAACCTGTGTGGGTGAAG 1560
DB      1572 AACTTTGGGAAAGTGTTCAGCGGACGCTGCGAGCCGACAAACCTGTGTGGGTGAAG 1631
QY      1561 TCTTGTGAGAGAGCTCCCACTGACCTCAAGGCAAGTTTTCACAGAAAGGAGATC 1620
DB      1632 TCTTGTGAGAGAGCTCCCACTGACCTCAAGGCAAGTTTTCACAGAAAGGAGATC 1691
QY      1621 CTGAAGCAGTACAGCCACCCCAACATGTGTGTCTGATTTGTGTGCAACCAAGAGCAG 1680
DB      1692 CTGAAGCAGTACAGCCACCCCAACATGTGTGTCTGATTTGTGTGCAACCAAGAGCAG 1751
QY      1681 CCCATCTTACATGATGAGAGCTTGTGACAGGGGGGAGACTTCTGTACCTTCCGACG 1740
DB      1752 CCCATCTTACATGATGAGAGCTTGTGACAGGGGGGAGACTTCTGTACCTTCCGACG 1811
QY      1741 GAGGGGCGCGCTGCGGAGTGAAGACTTGTGCTGACAGTGTGGGAGATGACGCTGTGC 1800
DB      1812 GAGGGGCGCGCTGCGGAGTGAAGACTTGTGCTGACAGTGTGGGAGATGACGCTGTGC 1871
QY      1801 ATGAGTACTCTGAGAGCAAGTGTCTATCCACCGGAGACTGTGCTCTGGAATCTGCTG 1860
DB      1872 ATGAGTACTCTGAGAGCAAGTGTCTATCCACCGGAGACTGTGCTCTGGAATCTGCTG 1931
QY      1861 GTGACAGAGAAAGATCTCTGAAGATCAGTGAATTTGGGATGTCTCCGAGAGAAAGCCGAT 1920

```

Db 1932 GTACAGAGAGAAATGCTCTGAAGATCAGTGAATCTTGGGATGTCCTCCAGAGAGAGCCGAT 1991
Qy 1921 GGGGCTATATCAGAGCTTGAAGGGGGCTTGAACAAAGTCCCGTGAAGTGAAGCCGACTGAG 1980
Db 1992 GGGGCTATATCAGAGCTTGAAGGGGGCTTGAACAAAGTCCCGTGAAGTGAAGCCGACTGAG 2051
Qy 1981 GGGCTTAACATACAGGGCGCTACTCTCCGAAAGGAGAGTGTGAAGCTTTGGCATCTTGCTC 2040
Db 2052 GGGCTTAACATACAGGGCGCTACTCTCCGAAAGGAGAGTGTGAAGCTTTGGCATCTTGCTC 2111
Qy 2041 TGGAGAGCCTTACAGCTTGAAGGGGGCTTCCCGTATCCCACTCAAGCATCAGACAGACAG 2100
Db 2112 TGGAGAGCCTTACAGCTTGAAGGGGGCTTCCCGTATCCCACTCAAGCATCAGACAGACAG 2171
Qy 2101 GAGTTGTGAAGAGAGGGGGGGCTTCCCGTATCCCACTCAAGCATCAGACAGACAG 2160
Db 2172 GAGTTGTGAAGAGAGGGGGGGCTTCCCGTATCCCACTCAAGCATCAGACAGACAG 2231
Qy 2161 AGGCTCATGAGAGAGTGTGGGCTTATGAGCTTGAAGAGGGGGCTTCAAGCATCAGACATC 2220
Db 2232 AGGCTCATGAGAGAGTGTGGGCTTATGAGCTTGAAGAGGGGGCTTCAAGCATCAGACATC 2291
Qy 2221 TACCAAGAGCTGCAAGAGCATCCGAAAGCGGCATCGG 2256
Db 2292 TACCAAGAGCTGCAAGAGCATCCGAAAGCGGCATCGG 2327

RESULT 3

US-10-240-965-256
Sequence 256, Application US/10240965
Publication No. US20030165924A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: SHIFFMAN, Dov
APPLICANT: SOMOGYI, Roland
APPLICANT: LAMN, Richard M.
APPLICANT: SEILHAMER, Jeffrey J.
APPLICANT: PORTER, Gordon J.
APPLICANT: MIKITA, Thomas
APPLICANT: TAI, Julie
TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
FILE REFERENCE: PA-0025 PCT
CURRENT APPLICATION NUMBER: US/10/240,965
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: 60/195,106
PRIOR FILING DATE: 2000-04-05
NUMBER OF SEQ ID NOS: 276
SOFTWARE: PERL Program
SEQ ID NO 256
LENGTH: 2889
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20030165924A1 997347.6
US-10-240-965-256

Query Match 89.7%; Score 2023.4; DB 16; Length 2889;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 225; Conservative 0; Mismatches 1; Indels 21; Gaps 2;

Qy 1 ATGGGCTTCTCTTGTGAGCTGTGAGCCCGAGGGGCTTGAAGCAAAATCAG 60
Db 188 ATGGGCTTCTCTTGTGAGCTGTGAGCCCGAGGGGCTTGAAGCAAAATCAG 247
Qy 61 GAGGCCAGCTTCTCTTGTGAGAGGAGCATGAAGAGTGAAGGCCAGCGGTCAGAGT 120
Db 248 GAGGCCAGCTTCTCTTGTGAGAGGAGCATGAAGAGTGAAGGCCAGCGGTCAGAGT 307
Qy 121 GACAGGAGATGAGAGAGCTCTTCAACCAATGTCCTTGAAGAGAGTGGGGCCAGAGC 180
Db 308 GACAGGAGATGAGAGAGCTCTTCAACCAATGTCCTTGAAGAGAGTGGGGCCAGAGC 367

Qy 181 CCGGCCATCAGCCCTTGAAGAGCCCATCAGTCACTCTGGGCTGAAGTCAACAGCCAACT 240
Db 368 CCGGCCATCAGCCCTTGAAGAGCCCATCAGTCACTCTGGGCTGAAGTCAACAGCCAACT 427
Qy 241 GA-GGGCTTGAAGCCGCTTCTGAGGAGAGCCGAGAGATCTGAATCTCAGGGCCCTGAG 299
Db 428 GAGGGCTTGAAGCCGCTTCTGAGGAGAGCCGAGAGATCTGAATCTCAGGGCCCTGAG 487
Qy 300 CAACTGAGCTTCTTCACTCCGGAAGGAGAGAGTCTTCCCAAGACTTAAGAGAGAG 359
Db 488 CAACTGAGCTTCTTCACTCCGGAAGGAGAGAGTCTTCCCAAGACTTAAGAGAGAG 547
Qy 360 GAGAGCTGAGAGAGAGAGTCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
Db 548 GAGAGCTGAGAGAGAGAGTCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 607
Qy 420 CCAATACGAGCTTGGAG 479
Db 608 CCAATACGAGCTTGGAG 667
Qy 480 CAAAG 539
Db 668 CAAAG 727
Qy 540 TGCTCAACCAACCGCTATGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 599
Db 728 TGCTCAACCAACCGCTATGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 787
Qy 600 CCAACAGCTTCTGAG 659
Db 788 CCAACAGCTTCTGAG 847
Qy 660 TTGCATCTGAAGAGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 719
Db 848 TTGCATCTGAAGAGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 907
Qy 720 GGTGTGAGCATTCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779
Db 908 GGTGTGAGCATTCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 967
Qy 780 GTACCAAGGCTTCTGAG 839
Db 968 GTACCAAGGCTTCTGAG 1027
Qy 840 CCAATGAGTCACTGAG 899
Db 1028 CCAATGAGTCACTGAG 1087
Qy 900 GCTGACTGTGAG 959
Db 1088 GCTGACTGTGAG 1147
Qy 960 CACCGAGATGTGTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1019
Db 1148 CACCGAGATGTGTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1207
Qy 1020 TGAAG 1079
Db 1208 TGAAG 1267
Qy 1080 GCAAG 1139
Db 1268 GCAAG 1327
Qy 1140 GCAAG 1199
Db 1328 GCAAG 1387
Qy 1200 GCTCTGAG 1259
Db 1388 GCTCTGAG 1447

1260 GACACCCAGCTGAGATCTCTTAAGGCCATCTCAGAAATCTTCCGCCCAAGTTCTC 1319
1448 GACACCCAGCTGAGATCTCTTAAGGCCATCTCAGAAATCTTCCGCCCAAGTTCTC 1507
1320 G----- 1320
1508 GCTCCCTCCACCGCTGAGCTCATTCGAGAGTGCAGAAACCCCTGATGAGAGCTGTG 1567
1321 ----- 1320
1568 GTACCAAGGAGCCATCCGAGGCGAGAGTGTGCTGAGCTGTGAGCACTTGGGAGCTT 1627
1321 ----- 1320
1628 CTTGTGCGGAGAGCCAGGGCAAGAGATAGTGTGTGCTGTGCTGTGAGATGCTT 1687
1321 ----- 1320
1688 GCGCCGCGCACTTATCATCCAGTCTTGTGATTAACCTGTACCGATGGAGGGAGGCTT 1747
1350 TCTTAGCATCTTTTGTCTATGACCACTTATGAGCAGCCAGCAGCCCTTACCAAGAA 1409
1748 TCTTACCATCTTGTCTCATCCAGCACTTATGAGCAGCCAGCAGCCCTTACCAAGAA 1807
1410 GAGTGTGTGTCTGTGACAGGGCTGTGCGCAAGCAAGTGGGTGTGTAACATAGAGA 1469
1808 GAGTGTGTGTCTGTGACAGGGCTGTGCGCAAGCAAGTGGGTGTGTAACATAGAGA 1867
1470 CTTGTGTGTGTGAGCAGATTGAGCAGGGGAACTTTGCGAAGTGTTCAGCGGAGCTT 1529
1868 CTTGTGTGTGTGAGCAGATTGAGCAGGGGAACTTTGCGAAGTGTTCAGCGGAGCTT 1927
1530 GCGAGCCGACACACCCTGTGTGCGGTGAAGTCTTGTGAGAGACGCTTCCACCTGACT 1589
1928 GCGAGCCGACACACCCTGTGTGCGGTGAAGTCTTGTGAGAGACGCTTCCACCTGACT 1987
1590 CAAGGCGAAGTTCTTACAGGAAGCGAGATCTTGAAGCAATTAAGCAACCCCAACATCT 1649
1988 CAAGGCGAAGTTCTTACAGGAAGCGAGATCTTGAAGCAATTAAGCAACCCCAACATCT 2047
1650 GCGCTCATTTGTGTGTGACAGCAAGAGAGAGCCCATCTTACATGTCATGAGCTTGTGCA 1709
2048 GCGCTCATTTGTGTGTGACAGCAAGAGAGAGCCCATCTTACATGTCATGAGCTTGTGCA 2107
1710 GGGGGGCGACTTCTTGACTTCTTCCGACGAGAGGGGCGCGCTGTGGGTGAAGACTCT 1769
2108 GGGGGGCGACTTCTTGACTTCTTCCGACGAGAGGGGCGCGCTGTGGGTGAAGACTCT 2167
1770 GCTGCAATGTGTGGGAGTGTGACAGCTGTGTGAGTGTGAGTGTGAGAGCAAGTGTGACT 1829
2168 GCTGCAATGTGTGGGAGTGTGACAGCTGTGTGAGTGTGAGTGTGAGAGCAAGTGTGACT 2227
1830 CCACCGGAGACTTGTGTGACAGCAAGAGAGAGCCCATCTTACATGTCATGAGCTTGTGCA 1889
2228 CCACCGGAGACTTGTGTGACAGCAAGAGAGAGCCCATCTTACATGTCATGAGCTTGTGCA 2287
1890 TGACTTTGGGATGTCCGAGAGAGAGCCGATGGGGTCTATGCAAGCTTCAAGGGGCGCTCAG 1949
2288 TGACTTTGGGATGTCCGAGAGAGAGCCGATGGGGTCTATGCAAGCTTCAAGGGGCGCTCAG 2347
1950 ACNAAGTCTCCGCTGAGAGTGTGACCGGACCTTGAAGGCTTAACTAAGGCGCTTCTCCGA 2009
2348 ACNAAGTCTCCGCTGAGAGTGTGACCGGACCTTGAAGGCTTAACTAAGGCGCTTCTCCGA 2407
2010 AACGCAAGTGTGAGCTTTGGGATCTTGTCTGTGGAGAGACTTCAAGCTTGGGGGCGCTTCCC 2069
2408 AACGCAAGTGTGAGCTTTGGGATCTTGTCTGTGGAGAGACTTCAAGCTTGGGGGCGCTTCCC 2467
2070 CTATCCCAACTGAGCAATCAGACAGACAGGAGTTTGTGAGAGAGGGGGGCGCTTGTGCC 2129
2468 CTATCCCAACTGAGCAATCAGACAGACAGGAGTTTGTGAGAGAGGGGGGCGCTTGTGCC 2527
2130 CTGCCAGAGCTGTGTCTGTATGCTGCGGTTCAGGCTCATGAGAGAGTGTGGGCGCTATGA 2189

2528 CTGCCAGAGCTGTGTCTGTATGCTGCGGTTCAGGCTCATGAGAGAGTGTGGGCGCTATGA 2587
2190 GCGTGGGAGAGCGGCGGAGCTTACAGCAACATTTACAGAGAGCTGTGAGAGATCCGAAGCG 2249
2588 GCGTGGGAGAGCGGCGGAGCTTACAGCAACATTTACAGAGAGCTGTGAGAGATCCGAAGCG 2647
2250 GCATCGG 2256
2648 GCATCGG 2654

RESULT 4

US-09-918-995-1503
; Sequence 1503, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918, 995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235, 076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FASTSeq for Windows Version 3.0
; SEQ ID NO 1503
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(449)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-1503

Query Match

Best Local Similarity 99.4%; Pred. No. 3, 9e-72; Matches 328; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1707 GCAGGGGGGCGACTTCTTCTGACCTTCTCCGACAGAGGGGGGCGCGCTGTGGGTGAAGAC 1766
120 GCGAGGGGGGCGACTTCTTCTGACCTTCTCCGACAGAGGGGGGCGCGCTGTGGGTGAAGAC 179
1767 TCTGCTGAGATGTGTGGGAGTGTGACAGCTGTGTGAGTGTGAGAGCAAGTGTG 1826
180 TCTGCTGAGATGTGTGGGAGTGTGACAGCTGTGTGAGTGTGAGAGCAAGTGTG 239
1827 CATCCACCGGAGACTTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1886
240 CATCCACCGGAGACTTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 239
1887 CAGTGACTTTGGGATGTCCGAGAGAGAGCCGATGGGGTCTATGCAAGCTTCAAGGGGCGCT 1946
300 CAGTGACTTTGGGATGTCCGAGAGAGAGCCGATGGGGTCTATGCAAGCTTCAAGGGGCGCT 359
1947 CAGACAGTCTCCGCTGAGAGTGTGACCGGACCTTGAAGGCTTAACTAAGGCGCTTCTC 2006
360 CAGACAGTCTCCGCTGAGAGTGTGACCGGACCTTGAAGGCTTAACTAAGGCGCTTCTC 419
2007 CGAAGCGAGCTGTGAGCTTTGGGATCTT 2036
420 CGAAGCGAGCTGTGAGCTTTGGGATCTT 449

RESULT 5

US-09-948-802-5
; Sequence 5, Application US/09948802
; Publication No. US20020172981A1
; GENERAL INFORMATION:
; APPLICANT: ROBINSON, KEITH E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND

TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
FILE REFERENCE: NMI-090
CURRENT APPLICATION NUMBER: US/09/948,802
CURRENT FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: 09/387,212
PRIOR FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 361
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-09-948-802-5

Query Match 13.1%; Score 296.2; DB 9; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.9e-64;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

QY 1613 CGAGATCTCGAAGCAGTACAGCCCAACATCGTCTCATTTGTCGTGACCC 1672
DB 20 CNAAGATCTCGAAGCAGTACAGCCCAACATCGTCTCATTTGTCGTGACCC 79
QY 1673 AGAAGCAGCCCATCTACATGTCATGAGCTTGTGACAGGGGGGCACTTCTGACCTTCC 1732
DB 80 AGAAGCAGCCCATCTACATGTCATGAGCTTGTGACAGGGGGGCACTTCTGACCTTCC 139
QY 1733 TCCGACAGGAGGGGGCCCGCTCGGGGTGAAGACTTGTGTCAGATGTTGGGGATGACAG 1792
DB 140 TCCGACAGGAGGGGGCCCGCTCGGGGTGAAGACTTGTGTCAGATGTTGGGGATGACAG 199
QY 1793 CTCTGGCATGAGTACTGAGAGCAAGTGTGATCCACCGGGGACCTGCTGTCCGA 1852
DB 200 CTCTGGCATGAGTACTGAGAGCAAGTGTGATCCACCGGGGACCTGCTGTCCGA 259
QY 1853 ACTGCTGTGACAGAGAAATGTCCTGAAG-ATCAGTACTTT-GGATGTCCGAGA 1910
DB 260 ACTGCTGTGACAGAGAAATGTCCTGAAGATCAGTACTTTGGGGATGTCCGAGA 319
QY 1911 GGAAGCC--GATGGGTCTATGACGCTCAAGGGGGCTTCAAG 1949
DB 320 GGAAGCCGATTGGGGTCTATGACGCTCAAGGGGGCTTCAAG 361

RESULT 6
US-10-121-925-5
Sequence 5, Application US/10121925
Publication No. US20030104505A1
GENERAL INFORMATION:
APPLICANT: ROBINSON, KEITH E.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
FILE REFERENCE: NMI-090
CURRENT APPLICATION NUMBER: US/10/121,925
CURRENT FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: US/09/948,802
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: 09/387,212
PRIOR FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 361
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-10-121-925-5

Query Match 13.1%; Score 296.2; DB 15; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.9e-64;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

QY 1613 CGAGATCTCGAAGCAGTACAGCCCAACATCGTCTCATTTGTCGTGACCC 1672
DB 20 CNAAGATCTCGAAGCAGTACAGCCCAACATCGTCTCATTTGTCGTGACCC 79
QY 1673 AGAAGCAGCCCATCTACATGTCATGAGCTTGTGACAGGGGGGCACTTCTGACCTTCC 1732
DB 80 AGAAGCAGCCCATCTACATGTCATGAGCTTGTGACAGGGGGGCACTTCTGACCTTCC 139
QY 1733 TCCGACAGGAGGGGGCCCGCTCGGGGTGAAGACTTGTGTCAGATGTTGGGGATGACAG 1792
DB 140 TCCGACAGGAGGGGGCCCGCTCGGGGTGAAGACTTGTGTCAGATGTTGGGGATGACAG 199
QY 1793 CTCTGGCATGAGTACTGAGAGCAAGTGTGATCCACCGGGGACCTGCTGTCCGA 1852
DB 200 CTCTGGCATGAGTACTGAGAGCAAGTGTGATCCACCGGGGACCTGCTGTCCGA 259
QY 1853 ACTGCTGTGACAGAGAAATGTCCTGAAG-ATCAGTACTTT-GGATGTCCGAGA 1910
DB 260 ACTGCTGTGACAGAGAAATGTCCTGAAGATCAGTACTTTGGGGATGTCCGAGA 319
QY 1911 GGAAGCC--GATGGGTCTATGACGCTCAAGGGGGCTTCAAG 1949
DB 320 GGAAGCCGATTGGGGTCTATGACGCTCAAGGGGGCTTCAAG 361

RESULT 7
US-10-280-576-19
Sequence 19, Application US/10280576
Publication No. US20040044405A1
GENERAL INFORMATION:
APPLICANT: WOLFE, Matthew R.
TITLE OF INVENTION: VASCULAR STENT OR GRAFT COATED OR IMPREGNATED WITH PROTEIN
FILE REFERENCE: 09820.189
CURRENT APPLICATION NUMBER: US/10/280,576
CURRENT FILING DATE: 2002-10-22
PRIOR APPLICATION NUMBER: 60/343,732
PRIOR FILING DATE: 2001-10-25
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.1
SEQ ID NO 19
LENGTH: 1779
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-10-280-576-19

Query Match 12.0%; Score 271.6; DB 17; Length 1779;
Best Local Similarity 57.0%; Pred. No. 3.6e-58;
Matches 518; Conservative 0; Mismatches 384; Indels 6; Gaps 1;

QY 1327 TACCGATGGAAGGGGAGGCTTCTTACGATCTTGTCTCATGACCACTACTGAGC 1386
DB 844 TTCCGATTGAGGGACCAACATTTGCGAGATCTAGAGATGATCATGATCATGATCAG 903
QY 1387 ACCCAGACCCCTCCACCAAGAAAGTGTGTTGCTCTGACAGGGCTGTGCCAAGAC 1446
DB 904 TCCGATTGCGAGTACCGTGAATCGGGAGCATCTCGACGACCCGTTTCCGGAG 963
QY 1447 AAGTGGTGTGACCATGAGACCTGTGTGGGTGAGCAGATTTGAGCGGGGAATTT 1506
DB 964 CGCTGGAGCTGTGACCAAGATGATGTGTGATCTTGTGAAGGATTTGTGGGGAATTT 1023
QY 1507 GCGGAATGTTCAGCGGAGCGCTGCGAGCGACCAACCCCTGTGCGGGTGAAGTCTTGT 1566
DB 1024 GGGGATGTCTACAGGCCAACTGAAGTCCACCACTGATGTGCTGTCAAACTGT 1083
QY 1567 CGAGAGCGCTCCCACTGACCTCAAGGCCAATTTCTACAGGAAGCAGAGATCTGAAG 1626
DB 1084 CGAATACCTGCGCGACGAGAGAGCGTAATTTCTACAGGAAGCGGCGATCTCAAG 1143
QY 1627 CAGTACGACCCCAACATCTGTGCTCTCATTTGTCGTGACCCCAAGAGAGCCCATC 1686
DB 1144 CAAATGATCATCTCAATATCTGTAAATGATTTGTGATTTGTGTCAGAGAGCCCATC 1203

QY 1687 TACATGTCATGAGAGCTTGTGCAAGGGGCGAATTCTTGAACCTTCTCCGCAAGAGGGG 1746
DB 1204 ATGATTTGTATGAAATTTGGTGTCTGGTGTCTGCTTTTAACTTATTAGCAAGAACTCC 1263
QY 1747 GCCCGCTCGGGGTGAAGACTGTCTGCAGATGTGGGGATGACCTGTGCATGAG 1806
DB 1264 AATGCGCTCAACCACTGCCAACAATGGGCAATGTGCAGAGATCGGGCGGAGGATGCCA 1323
QY 1807 TACCTGAGAGCAAGTGTGCATCCACCGGACCTGTGCTGCAATGCTGTGACA 1866
DB 1324 TATCTGAGATCCAAAATGCAATTCAATCGGATCTGGCGGCGGTATTTGTCTGTGAC 1383
QY 1867 GAGAAAGATGCTGTAAGATCACTGACTTTGGGATCTCCGAGAGAGACCGATGGGGTTC 1926
DB 1384 TTGGACACAGTGTGAAGATCTCCGATTTCCGAATGTCTCGCAGGAA-----GAGGAA 1437
QY 1927 TATGACGCTCAAGGGGGGCTCAGACAAAGTCCCGTGAAGTGAACCGACCTGAGGCCCTT 1986
DB 1438 TATATGTTTCCGATGCGATGAAACAAATACCTGTGAAGTGAACAGCTCCGAGGCTTGG 1497
QY 1987 AACTACGCGCGCTACTCTCCGAAAGCGAGTGTGAGCTTTGGCATTTGTCTTGGAG 2046
DB 1498 AATTTCGGCAAGTACCTTGTGTGCGATGTGTGCTTATGTCATGATGATGTGGAG 1557
QY 2047 ACCTTACGCTGGGGGCTCTCCCTTATCCCACTCAGCAATACAGACAGCGAGGTTT 2106
DB 1558 ATCTTCTCCAAAGGGCCACACACCTTACTCGGCAATACCACTCAGAGCAAGAGCGC 1617
QY 2107 GTGAGAGAGGGGGGCGTCTGCGCCCTGCGAGCTGTGTCTGTATGCTCGGTTCAGGCTC 2166
DB 1618 ATCGATACGGGATATGTATGTCACAGCGGAGAGACGCGCCGAGATGTATCCGACTG 1677
QY 2167 ATGAGACAGTGTGGGCTTATGAGCCTGTGGCAGCGGCCCACTTACAGCAATCTACCA 2226
DB 1678 ATGCTTCAGTGTGGGCGAGCGCGCAATCCGACCGCATTTGATGATGATCTCAAT 1737
QY 2227 GAGCTGCA 2234
DB 1738 GTGGTGA 1745

RESULT 8

US-10-003-295-3
; Sequence 3, Application US/10003295
; Publication No. US20020168741A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001183DIV
; CURRENT APPLICATION NUMBER: US/10/003,295
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-003-295-3

Query Match 10.1%; Score 227.2; DB 13; Length 15297;
Best Local Similarity 71.1%; Pred. No. 7.3e-47;
Matches 394; Conservative 0; Mismatches 8; Indels 152; Gaps 1;

QY 1 ATGGGCTTCTCTTCTAGCTGTGACGCCCGCAGGCGCAAGGAGTCTTGCAGCAAAATGCA 60
DB 2563 ATGGGCTTCTCTTCTAGCTGTGACGCCCGCAGGCGCAAGGAGTCTTGCAGCAAAATGCA 2622
QY 61 GAGGCGAGCTTCTGCTACTGAGGCGATGAGAAAGTGTGAGCGGCTCAAGAGT 120
DB 2623 GAGGCGAGCTTCTGCTACTGAGGCGATGAGAAAGTGTGAGCGGCTCAAGAGT 2682

QY 121 GACAGGAGTATGACAGAGTGTGACCAATGCTCTGACAGAGACAGTGGGGCCAGAGC 180
DB 2683 GACAGGAGTATGACAGAGTGTGACCAATGCTCTGACAGAGACAGTGGGGCCAGAGC 2742
QY 181 CGGGCCATGAGCCCTGACAGGCCCATCAGT----- 211
DB 2743 CGGGCCATGAGCCCTGACAGGCCCATCAGTGGGCTCTATAGGACTCTGTGGGT 2802
QY 212 ----- 211
DB 2803 GCTGGGTATGCTCTTCTCTCTCTGCGGGGCGCTTGGGAGTGTGCTGAGAT 2862
QY 212 ----- 211
DB 2863 CTGGAGCGCAATGCTTGGAGCAATGTGCCCCCTCTGCTCCCATCTGTGCTG 2922
QY 212 ---AGTCTGGGCTGATCATCAGCAAACTGAGGGCTGAGCGCTTGTGCGGACG 268
DB 2923 TATAGTCTGGGCTGAGATCATCAGCAAACTGAGGGCTGAGCGCTTGTGCGGACG 2982
QY 269 ACGCAGAGATCTGAATCTGAGGGCGCTGAGCAAGCTGAGCTGTCTATTCGGGAACTGC 328
DB 2983 ACGCAGAGATCTGAATCTGAGGGCGCTGAGCAAGCTGAGCTGTCTATTCGGGAACTGC 3042
QY 329 AGCAGCTTGGCAAGCTTACAGGAGCTGTGAGAGTGTGACAGAGAGCTTACCAAGA 388
DB 3043 ACGAGCTTGGCAAGCTTACAGGAGCTGTGAGAGTGTGACAGAGAGCTTACCAAGA 3102
QY 389 CCACAGCCAGGAG 402
DB 3103 TGAGCGGCGCAGC 3116

RESULT 9

US-10-660-763-3
; Sequence 3, Application US/10660763
; Publication No. US20040063130A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001183DIVIT
; CURRENT APPLICATION NUMBER: US/10/660,763
; CURRENT FILING DATE: 2003-09-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-660-763-3

Query Match 10.1%; Score 227.2; DB 17; Length 15297;
Best Local Similarity 71.1%; Pred. No. 7.3e-47;
Matches 394; Conservative 0; Mismatches 8; Indels 152; Gaps 1;

QY 1 ATGGGCTTCTCTTCTAGCTGTGACGCCCGCAGGCGCAAGGAGTCTTGCAGCAAAATGCA 60
DB 2563 ATGGGCTTCTCTTCTAGCTGTGACGCCCGCAGGCGCAAGGAGTCTTGCAGCAAAATGCA 2622
QY 61 GAGGCGAGCTTCTGCTACTGAGGCGATGAGAAAGTGTGAGCGGCTCAAGAGT 120
DB 2623 GAGGCGAGCTTCTGCTACTGAGGCGATGAGAAAGTGTGAGCGGCTCAAGAGT 2682
QY 121 GACAGGAGTATGACAGAGTGTGACCAATGCTCTGACAGAGACAGTGGGGCCAGAGC 180
DB 2683 GACAGGAGTATGACAGAGTGTGACCAATGCTCTGACAGAGACAGTGGGGCCAGAGC 2742
QY 181 CGGGCCATGAGCCCTGACAGGCCCATCAGT----- 211
DB 2743 CGGGCCATGAGCCCTGACAGGCCCATCAGTGGGCTCTATAGGACTCTGTGGGT 2802

Qy	212	-----	211
Db	2803	GCTGGCGTATCTGCCTTCTCCTTCTCTCTGAGGGGCCCTCTGGGGCAGTGGCTGGAGAT	2862
Qy	212	-----	211
Db	2863	CTGGCAGGCCAATGCTTGGGAGCCATTGTGCCCCCTCCCTGCTCCCTCCCACTGTGTGTG	2922
Qy	212	---AGTCCTGGGCTGAGATCAACAGCCAACTGAAGGCTTGAGCCGCTTGCTGCGGACG	268
Db	2923	TATAGTCCTGGGCTGAGATCAACAGCCAACTGAAGGCTTGAGCCGCTTGCTGCGGACG	2982
Qy	269	ACGCAGAGGATCTGAATCTCAGGGCCCCCTGAGCAAGCTAGCTGTCTATCCGGGAAACGGC	328
Db	2983	ACGCAGAGGATCTGAATCTCAGGGCCCCCTGAGCAAGCTAGCTGTCTATCCGGGAAACGGC	3042
Qy	329	AGCAGCTTCGAAAGACTCACAGCGAGCAGTGGCAGCACTGACAGGAGAGCTCACCAAGA	388
Db	3043	AGCAGCTTCGAAAGACTCACAGCGAGCAGTGGCAGCACTGACAGGAGAGCTCACCAAGA	3102
Qy	389	CCACAGCCAGAC	402
Db	3103	TGAGCGGCACAC	3116

Query	Match	Similarity	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
US-10-101-510-525	Sequence 525: Application US/10101510									
US-10-101-510-525	Publication No. US20030148295A1									
US-10-101-510-525	GENERAL INFORMATION:									
US-10-101-510-525	APPLICANT: WAN, JACKSON									
US-10-101-510-525	APPLICANT: WAN, YIXIN									
US-10-101-510-525	TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE									
US-10-101-510-525	FILE REFERENCE: 15117.0012									
US-10-101-510-525	CURRENT APPLICATION NUMBER: US/10/101,510									
US-10-101-510-525	PRIOR FILING DATE: 2002-03-20									
US-10-101-510-525	PRIOR APPLICATION NUMBER: 60/276,947									
US-10-101-510-525	PRIOR FILING DATE: 2001-03-20									
US-10-101-510-525	NUMBER OF SEQ ID NOS: 805									
US-10-101-510-525	SOFTWARE: PatentIn Ver. 2.1									
US-10-101-510-525	SEQ ID NO 525									
US-10-101-510-525	LENGTH: 3875									
US-10-101-510-525	TYPE: DNA									
US-10-101-510-525	ORGANISM: Homo sapiens									
US-10-101-510-525	US-10-101-510-525									
Query Match	8.9%	Score 201.2;	DB 15;	Length 3875;						
Best Local Similarity	54.9%	Pred. No. 2e-40;								
Matches 423;	Conservative 0;	Mismatches 338;	Indels 9;	Gaps 1						

Db	2269	GGGAGGACCAAGCTGTCCTCCGCGGAGCTAATGGCCATGCTGACGAGGAGATAGCATCTGGCA	2328
Qy	1802	TGAGATACCTTGAGAGCAAGTGTCTGCATCCACCGGGACCTGGCTGCTCCGGAATCTGCTGG	1861
Db	2329	TGAACCTACCTCAGTAATCACAATTATGTGTCCACGGGAGCTGGCTGCCAGAAACATCTTTGG	2388
Qy	1862	TGACAGAGAAAGAAATGTCCTGAAAGATCAGTGAATTTGGGAATGTCCCGAGAGAAAGCCGATG	1921
Db	2389	TGAATCAAAACCTGTGCTGCAGAGGTGTCTGACTTTTGGCTGTGACTCGGCTCTCGATGACT	2448
Qy	1922	GGGTCTATGCAAGCTCAGGGGGGCTCAGACAAAGTCCCCCGTGAAGTGAACCGCAACTGAG	1981
Db	2449	TTGATGGCAATACGAAACCCAGGGAGAAAGATCCCTATTCGGTTGACAGCCCTCTTAAG	2508
Qy	1982	CCCTTAACCTACGGGCGCTACTCCTCCGAAAGCGACGTGTGAGCTTTGGCATCTTGGCT	2041
Db	2509	CCATTGGCCATCCGANTTTTACCAACAGCCAGCGATGTGTGAGGCTTTGGATTTGTGATGT	2568
Qy	2042	GGGAGACCTTCAGGCTTGGGGGCGCTCCCCGTATCCCACTCAGCAATCAGCAGACAGGG	2101
Db	2569	GGGAGGTGTGAGTTTGGGGAACAGGCTTATGGGGAGATGAGGAATACAGGAGTTTATGA	2628
Qy	2102	AGTTTGTGAGAAAGGGGGGCGGTCTGGCCCTGACCAGAGCTGTGTCTCATGATCCCGTGTCA	2161
Db	2629	AGACATTGAGATGGGTACCGGTGGCCCTCTGTGTGACTGCTTGGCCCTCTGTATG	2688
Qy	2162	GGCTCATGAGACAGTCTGGGCGCTTATGAGCTTGGGACAGCGGCCAGGCTTC	2211
Db	2689	AGCTCATGAAAGACTCTGGGGCATATACCGTGGCCGCCCGGCACACTTC	2738

```

RESULT 11
US-10-384-339C-1
: Sequence 1, Application US/10384339C
: Publication No. US20040175703A1
GENERAL INFORMATION:
APPLICANT: Kreutzler, Roland
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING EXPRESSION OF A TARGET GE
FILE REFERENCE: 20200/2002
CURRENT APPLICATION NUMBER: US/10/384,339C
CURRENT FILING DATE: 2003-03-07
PRIOR APPLICATION NUMBER: PCT/EP02/00152
PRIOR FILING DATE: 2002-01-09
PRIOR APPLICATION NUMBER: DE 10100586.5
PRIOR FILING DATE: 2001-01-09
PRIOR APPLICATION NUMBER: DE 10155280.7
PRIOR FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: DE 10158411.3
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: DE 10160151.4
PRIOR FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 173
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 2955
TYPE: DNA
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
TITLE: Eph A1
PATENT DOCUMENT NUMBER: NM00532
US-10-384-339C-1

Query Match      8.8%; Score 198; DB 18; Length 2955;
Best Local Similarity 54.7%; Pred. No. 1,3e-39;
Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

Oy      1451 GGGTGTCAACCATGATGAGCACTGTGTGGTGAGAGCAATTGACGGGGGAATTGGGCG 1510
Db      1874 GGGAGCTTGATTCACGCTGCTGATGTGTGACACTCTCATAGAGAAAGAGAGATTGGGG 1933

Oy      1511 AAGTGTTCAGGGAGCCCTTCAGGCTTCCCGACCGAGGACTTGGCCATTAAAGA 1561
Db      1934 AAGTGTATGAGGGACCTTCAGGCTTCCCGACCGAGGACTTGGCCATTAAAGA 1993

```

QY 1562 CTTGTGAGAGAGGCTCCCACTGACCTCAAGGCCAAGTTTCTTACAGAAAGCAGATCC 1621
DB 1994 CTTAAAGACATCCCAAGGAGGCGAGTGTGAACTTCTTCAAGAGCAATATCA 2053
QY 1622 TGAAGCAGTACAGCCCAACATCGTGTCTCAATTGGTGTCTGCAACCAAGAACAG 1681
DB 2054 TGGGCCAGTTTACCAACCCGCAATTTCTGATCTGGAAGGCGTGTCAACAAAGCAAAG 2113
QY 1682 CCATCTACATCGTCAATGAGCTTTGTGCAAGGGGGCGAATTCTGACCTTCTCCGACGG 1741
DB 2114 CGATCTATGATCAACAGAAATTTATGAGAAATCAACCCCTGATGCTTCTGAGGAGC 2173
QY 1742 AGGGGGCCCCCGGGGTGAAGACTGTGTGAGATGTGTGGGGAGATGCACTGTGCGCA 1801
DB 2174 GGAAGGACCAAGCTGTCTCCCTGGGCACTGAGGCAATGCTGCAAGGCGATATGCACTGGCA 2233
QY 1802 TGAAGTACCTGAGAGCAAGTGTGATCCACCGGGAACCTGCTGTGCGAACTGCTGG 1861
DB 2234 TGAATCACTCAAGTATATCAAAATTAATGTCCACCGGAACCTGCTGCGCAAAACATTTGG 2293
QY 1862 TGAAGAGAAAGATGTCTGAAAGATCAATGATCTTGGAGTGTCCGAGAGAAAGCCGATG 1921
DB 2294 TGAATCAAAACCTGTGCTGCAAGGTGTCTGACTTTGGCCTGACTCGGCTCTGATGATG 2353
QY 1922 GGGTCTATGACGCTCAGGGGGGCTCAGACAAAGTCCCCGTAAGTGAACCGCACTGAGG 1981
DB 2354 TGTATGCACTATGCAAAACCAAGAGAAAGATCCCTATCCGTTGAGACAGCCCTGGAAG 2413
QY 1982 CCCTTAACACGCGCGCTACTCTCCGAAAGGACGAGTGGAGCTTTGGCATTTGTCTCT 2041
DB 2414 CCATTCCTCATCGAATCTTTCACACAGCCAGCAATGTGTGAGCTTTGGATTTGTATGT 2473
QY 2042 GGGAGACCTTCAAGCTGCGGGGCTCCCTATCCCAACTCAGCAATCAGACAGACAGGG 2101
DB 2474 GGAAGGTGCTGAGCTTTGGGAGCAAGCTTATGGGGAGATGAGCAATCAGAGATTTATGA 2533
QY 2102 AGTTTGTGAGAAAGGGGGGCGCTGTCTGCTGCGCAAGCTGTCTGATGCCGTGTCA 2161
DB 2534 AGAGCATTTGAGATGGGTACCGGTTGCCCTCTGTGACATGCGCTGTGCGCTCTGTATG 2593
QY 2162 GGCTCATGAGACAGTGTGGGCTTATGAGCCTGGGAGCGGCGCCAGCTTC 2211
DB 2594 AGCTCATGAAGACTGCTGGGCAATATGACCTGTCCCGCGGCAACTTC 2643

RESULT 12

US-09-967-768A-144

Sequence 144, Application US/09967768A

Patent No. US20020150877A1

GENERAL INFORMATION:

APPLICANT: Augustus, Meena

TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu

FILE REFERENCE: 689290-72

CURRENT APPLICATION NUMBER: US/09/967,768A

CURRENT FILING DATE: 2001-09-28

PRIOR APPLICATION NUMBER: US/60/236,109

PRIOR FILING DATE: 2000-09-28

PRIOR APPLICATION NUMBER: US/60/236,034

PRIOR FILING DATE: 2000-09-28

PRIOR APPLICATION NUMBER: US/60/236,111

PRIOR FILING DATE: 2000-09-28

NUMBER OF SEQ ID NOS: 325

SOFTWARE: PatentIn version 3.0

SEQ ID NO 144

LENGTH: 3370

TYPE: DNA

ORGANISM: Homo sapiens

US-09-967-768A-144

Query Match 8.8%; Score 198; DB 9; Length 3370;
Best Local Similarity 54.7%; Pred. No. 1.3e-39;

Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;
QY 1451 GGGTCTGAACCATGAGACCTGTGTGTGGTGAGCAATTTGACCGGGAACTTTGGCG 1510
DB 1967 GGGAGCTTATCAGAGTGTGTGATGTGTGAGCACTGTCTATGAGAAAGAGAGTGTGGG 2026
QY 1511 AGTGTTCAGGGAGCGCTGCGAGCC-----GACAAACCCGTGTGGCGGTGAAGT 1561
DB 2027 AGTGTATGAGGAGACCTTCAGGCTCCCAAGCAGAGATGCAAGTGTGTGCAATTAAGA 2086
QY 1562 CTTGTGAGAGAGCGTCCCACTGACCTCAAGGCCAAGTTTCTTACAGAAAGCAGATCC 1621
DB 2087 CTTAAAGACATCCCAAGGAGGCGAGTGTGAACTTCTTCTGAGAGGCAATATCA 2146
QY 1622 TGAAGCAGTACAGCCCAACATGTGTCTCAATTGGTGTCTGCAACCAAGAACAG 1681
DB 2147 TGGGCCAGTTTACCAACCCGCAATTTCTGATCTGGAAGGCGTGTCAACAAAGCAAAG 2206
QY 1682 CCATCTACATCGTCAATGAGCTTTGTGCAAGGGGGCGAATTCTGACCTTCTCCGACGG 1741
DB 2207 CGATCTATGATCAACAGAAATTTATGAGAAATCAACCCCTGATGCTTCTGAGGAGC 2266
QY 1742 AGGGGGCCCCCGGGGTGAAGACTGTGTGAGATGTGTGGGGAGATGCACTGTGCGCA 1801
DB 2267 GGAAGGACCAAGCTGTCTCCCTGGGCACTGAGGCAATGCTGCAAGGCGATATGCACTGGCA 2326
QY 1802 TGAAGTACCTGAGAGCAAGTGTGATCCACCGGGAACCTGCTGTGCGAACTGCTGG 1861
DB 2327 TGAATCACTCAAGTATATCAAAATTAATGTCCACCGGGAACCTGCTGCGCAAAACATTTGG 2386
QY 1862 TGAAGAGAAAGATGTCTGAAAGATCAATGATCTTGGAGTGTGTCCGAGAGAAAGCCGATG 1921
DB 2387 TGAATCAAAACCTGTGCTGCAAGGTGTCTGACTTTGGCCTGACTCGCTCTGAGATGACT 2446
QY 1922 GGGTCTATGACGCTCAGGGGGGCTCAGACAAAGTCCCCGTAAGTGAAGCAGCACTGAGG 1981
DB 2447 TGTATGCACTATGCAAAACCAAGAGAAAGATCCCTATCCGTTGAGACAGCCCTGGAAG 2506
QY 1982 CCCTTAACACGCGCGCTACTCTCCGAAAGGACGAGTGGAGCTTTGGCATTTGTCTCT 2041
DB 2507 CCATTCCTCATCGAATCTTTCACACAGCCAGCAATGTGTGAGCTTTGGATTTGTATGT 2566
QY 2042 GGGAGACCTTCAAGCTGCGGGGCTCCCTATCCCAACTCAGCAATCAGACAGACAGGG 2101
DB 2567 GGAAGGTGCTGAGCTTTGGGAGCAAGCTTATGGGGAGATGAGCAATCAGAGATTTATGA 2626
QY 2102 AGTTTGTGAGAAAGGGGGGCGCTGTCTGCTGCGCAAGCTGTGTCTGATGCCGTGTCA 2161
DB 2627 AGAGCATTTGAGATGGGTACCGGTTGCCCTCTGTGACATGCGCTGTGCGCTCTGTATG 2686
QY 2162 GGCTCATGAGACAGTGTGGGCTTATGAGCCTGGGAGCGGCGCCAGCTTC 2211
DB 2687 AGCTCATGAAGACTGCTGGGCAATATGACCTGTCCCGCGGCAACTTC 2736

RESULT 13

US-10-354-358-101

Sequence 101, Application US/10354358

Publication No. US20030157082A1

GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc

APPLICANT: Hunter, John Joseph

APPLICANT: MacBeth, Kyle J.

APPLICANT: Tsai, Feng-Ying

APPLICANT: Lesoon, Andrea

APPLICANT: Lightcap, Eric S.

APPLICANT: Williamson, Mark

APPLICANT: Rudolph-Owen, Laura A.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING

TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,

TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,

TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,

TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,

TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
FILE REFERENCE: MPI02-020P1RMONIM
CURRENT APPLICATION NUMBER: US/10/354,358
PRIOR FILING DATE: 2003-01-30
PRIOR APPLICATION NUMBER: US 60/353,600
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: US 60/364,517
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/371,075
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: US 60/371,507
PRIOR FILING DATE: 2002-04-10
PRIOR APPLICATION NUMBER: US 60/372,984
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/374,194
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/382,995
PRIOR FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: US 60/385,023
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US 60/388,853
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: US 60/389,395
PRIOR FILING DATE: 2002-06-17
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 122
SOFTWARE: FaSTSeq for Windows Version 4.0
SEQ ID NO 101
LENGTH: 3370
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (94) ... (3048)
US-10-354-358-101

Query Match 8.8%; Score 198; DB 16; Length 3370;
Best Local Similarity 54.7%; Pred. No. 1.3e-39;

Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

QY 1451 GGGTGTGAACATGAGAGACCTGATGTTGGGTGAGCAGATTGACCGGGGAATTGGCG 1510
DB 1967 GGGAGCTTGAATCCAGGCTGCTGATGTGAGACTGTGATGAGAGAGAGAGATTGGGG 2026
QY 1511 AAGTGTTCAGCGGACCGCTTCGAGCC-----GACAAACCCCTGCTGGTGGTGAAT 1561
DB 2027 AAGTGTATCAGAGGAGACCTCAGGCTCCCGACGAGAGCTGCAAGACTGTGGCATTAGA 2086
QY 1562 CTGTGTGAGAGAGAGCTCCCACTGACTCAAGGCAAGTTTCAAGAGAGAGATCC 1621
DB 2087 CCTTAAAGACATATCCCAAGGTGGCCAGTGTGGAATCTTCTTCAAGAGGCAATATCA 2146
QY 1622 TGAAGCAGTACAGCAACCCCAACATGCTGCTCATTTGGTGTCTGACCCAGCAAGCAGC 1681
DB 2147 TGGGCCAGTTTACGACCCGCAATTTCTGATCTGGAAGGCGTGTCAAGAGGAAAGC 2206
QY 1682 CCATCTACATGCTCATGAGAGCTTGTGAGGGGGGCACTTCTGACCTTCTCCGACGG 1741
DB 2207 CGATCTATGATCATCAAGAAATTTATGAGATCAGCCCTGAGATGCTTCTGAGGGAGC 2266
QY 1742 AGGGGGCCCGCTGCGGGGTGAAGACTGTGCGAGATGTTGGGGGAATGCAAGCTGTGGCA 1801
DB 2267 GGGAGGACAGCTGTGCTCTGGGGAGCTAGTGGCCATGTGCAAGGGCAATGCAATGGCA 2326
QY 1802 TGAAGTACCTGAGAGCAAGTGTGATCCACCGGGAAGCTGTGCTCGAATGTGCTGG 1861
DB 2327 TGAATCACTTCAATGATCAAGAAATTTATGATCCAGGAGCCTGGTCCAGAAACATCTGG 2386
QY 1862 TGAAGAGAAATGTTCTGAAGATCAGTACTTTGGGATGTCCGAGAGAAAGCCGATG 1921
DB 2387 TGAATCAAACTGTGTGCAAGGTGTGACTTTGGCTGAGCTGCTCTGGAATGACT 2446

QY 1922 GGGTCTATCAGCTTCAGGGGCTTCAGACAAGTCCCGCTGAAGTGAACCGACCTGAGG 1981
DB 2447 TTGATGAGCACAATACGAACCCAGGAGAGAAATCCCTATCCGTTGAGACGCCCTGAAG 2506
QY 1982 CCTTAACTACGCGCGCTACTCTCCGAAAGGACCTGTGAGAGCTTTGGCATCTTGTCT 2041
DB 2507 CCATGGCCATCGAATCTTCAACACAGCGAGATGTGTGAGCTTTGGGAATGTGATGT 2566
QY 2042 GGGAGACCTTCAGCCCTGGGGGCTCCCTATCCCACTCAAGCATCAGACAGACGGG 2101
DB 2567 GGGAGGTGCTGAGCTTTGGGGACAGCTTATGGGAGATGACAAATCAGAGCTTATGA 2626
QY 2102 AGTTGTGAGAGAGGGGGGCGCTGTGCTGCGCCAGAGCTGTGTCTGATGCGGTGTCA 2161
DB 2627 AGAGCATTTAGAGATGTGTATCCGGTTGCCCTCTGTGAGACTGCCCTGCGCCCTGTATG 2686
QY 2162 GGCTCATGAGAGAGTGTGGGCTTATGAGCTTGGGAGCGGCCCACTTC 2211
DB 2687 AGCTCATGAAGAACTGCTGGGATATGACCGTCCGCGGCAACACTTC 2736

RESULT 14

US-10-210-120-19

/ Sequence 19, Application US/10210120

/ Publication No. US20030175736A1

/ GENERAL INFORMATION:

/ APPLICANT: Chinnaiyan, Arul M.

/ APPLICANT: Rubin, Mark A.

/ APPLICANT: Sreekumar, Arun

/ TITLE OF INVENTION: Expression Profile of Prostate Cancer

/ FILE REFERENCE: UM-07221

/ CURRENT APPLICATION NUMBER: US/10/210,120

/ PRIOR APPLICATION NUMBER: US 60/309,581

/ PRIOR FILING DATE: 2001-08-02

/ PRIOR APPLICATION NUMBER: US 60/334,468

/ PRIOR FILING DATE: 2001-11-15

/ NUMBER OF SEQ ID NOS: 123

/ SOFTWARE: PatentIn version 3.2

/ SEQ ID NO 19

/ LENGTH: 3370

/ TYPE: DNA

/ ORGANISM: Homo sapiens

US-10-210-120-19

Query Match 8.8%; Score 198; DB 16; Length 3370;
Best Local Similarity 54.7%; Pred. No. 1.3e-39;

Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

QY 1451 GGGTGTGAACATGAGAGACCTGATGTTGGGTGAGCAGATTGACCGGGGAATTGGCG 1510
DB 1967 GGGAGCTTGAATCCAGGCTGCTGATGTGAGACTGTGATGAGAGAGAGAGATTGGGG 2026
QY 1511 AAGTGTTCAGCGGACCGCTTCGAGCC-----GACAAACCCCTGCTGGTGGTGAAT 1561
DB 2027 AAGTGTATCAGAGGAGACCTCAGGCTCCCGACGAGAGCTGCAAGACTGTGGCATTAGA 2086
QY 1562 CTGTGTGAGAGAGAGCTCCCACTGACTCAAGGCAAGTTTCAAGAGAGAGAGATCC 1621
DB 2087 CCTTAAAGACATATCCCAAGGTGGCCAGTGTGGAATCTTCTTCAAGAGGCAATATCA 2146
QY 1622 TGAAGCAGTACAGCAACCCCAACATGCTGCTCATTTGGTGTCTGACCCAGCAAGCAGC 1681
DB 2147 TGGGCCAGTTTACGACCCGCAATTTCTGATCTGGAAGGCGTGTCAAGAGGAAAGC 2206
QY 1682 CCATCTACATGCTCATGAGAGCTTGTGAGGGGGGCACTTCTGACCTTCTCCGACGG 1741
DB 2207 CGATCTATGATCATCAAGAAATTTATGAGATCAGCCCTGAGATGCTTCTGAGGGAGC 2266
QY 1742 AGGGGGCCCGCTGCGGGGTGAAGACTGTGCGAGATGTTGGGGGAATGCAAGCTGTGGCA 1801
DB 2267 GGGAGGACAGCTGTGCTCTGGGGAGCTAGTGGCCATGTGCAAGGGCAATGCAATGGCA 2326


```

QY 1802 TGAAGTACCTGAGAGAGCAAGTGTGATCCACCGGAACTGTGCTGCTGCGAACTGCTCG 1861
DB 2327 TGAAGTACCTGAGATATCAATTAATATGTCACCGGAACTGTGCTGCGAACTGCTCG 2386
QY 1862 TGAAGTACCTGAGATATGCTGGAAGATGATGATGATGATGATGATGATGATGATGATG 1921
DB 2387 TGAATCAAAACCTGTGCTGCAAGGTGTCTGATCTTGATCTTGATCTTGATCTTGATCT 2446
QY 1922 GGGTCTATGACAGCTTCAGGGGGGCTTCAGCAAGTCCCGTGAAGTGAACCGGACCTGAG 1981
DB 2447 TTGATGAGCAATACGAAACCCAGGGAGAAAGATCCCTATCCGTTGAGACAGCCCTGAG 2506
QY 1982 CCCTTAATACGAGCGGCTACTCTCCGAAAGCACTGTGAGCTTTGGCATCTTGCTCT 2041
DB 2507 CCATTGCCATCGGATCTTACCAACGACGCGATGTGAGCTTTGGGATTTGATGT 2566
QY 2042 GGGAGACCTTCAGCGCGGGGGCTCCCGCTATCCAACTCAGCAATCAGACAGACGAG 2101
DB 2567 GGGAGGCTGCTGAGCTTTGGGGAGACAGCTTATGAGGAGATGAGCAATCAGAGGTTATGA 2626
QY 2102 AGTTGTGAGAGAGGGGGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2161
DB 2627 AAGAGCATTTAGAGATGGGTATCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2686
QY 2162 GGCTCATGAGAGATGCTGGGCTTATGAGCTTGGGAGCGGCGCCAGCTTC 2211
DB 2687 AGCTCATGAGAGATGCTGGGCTTATGAGCTTGGGAGCGGCGCCAGCTTC 2736

```

RESULT 15

```

US-10-776-827-82
/ Sequence 82, Application US/10776827
/ Publication No. US20040132086A1
/ GENERAL INFORMATION:
/ APPLICANT: Horwiltz, Jennifer
/ APPLICANT: Richer, Jennifer
/ TITLE OF INVENTION: Progestosterone Receptor-Regulated Gene Expression and Methods Relat
/ FILE OF INVENTION: Therefo
/ FILE REFERENCE: 2848-39
/ CURRENT APPLICATION NUMBER: US/10/776,827
/ CURRENT FILING DATE: 2004-02-10
/ PRIOR APPLICATION NUMBER: US/09/814,915
/ PRIOR FILING DATE: 2002-03-21
/ PRIOR APPLICATION NUMBER: 60/214,870
/ PRIOR FILING DATE: 2000-06-28
/ NUMBER OF SEQ ID NOS: 108
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 82
/ LENGTH: 3370
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-776-827-82

```

Query Match 8.8%; Score 198; DB 18; Length 3370;

Best Local Similarity 54.7%; Pred. No. 1.3e-39;

Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

```

QY 1451 GGGTGTGTAACCATGAGACCTGTGTGGGTGAGCAATTGACGAGGGAATTTGGCG 1510
DB 1967 GGGAGCTTATTCAGGCTGTGATGTGATGACATCTGTATGAGAGAGAGATTGGGG 2026
QY 1511 AAGTGTTCAGCGGACCTGCGAGCC-----GACAAACACCTGTGGCGGTGAAGT 1561
DB 2027 AAGTGTATGAGGAGACCTTCAGGCTCCGACGAGAGCTGCAAGACTGTGGCCATTAAAG 2086
QY 1562 CTTGTGAGAGAGAGCTCCCACTGACCTCAAGGCCAAGTTTCTACAGAGAGAGATCC 1621
DB 2087 CCTTAAAGACACATCTCCGAGGTGGCCAGTGTGAACTTCTTGAAGAGCAACTATCA 2146
QY 1622 TGAAGCAGTACAGCAACCCCAACATCGTGTGCTCATTTGGTGTCTGACCCAGAGCAGC 1681
DB 2147 TGGGCGAGTTTACGCCACCGCATATTTCTGATCTGAGAGGCGTGTCTCAAAAGCGAAGC 2206

```

```

QY 1682 CCATCTACATGCTCATGAGCTTTGCAAGGGGGGAGCTTCTCTGACCTTCTCCGACGG 1741
DB 2207 CCATCATGATATCATCAGAAATTTATGAGAAATGACAGCCCTGATGCTTTCTGAGGGAGC 2266
QY 1742 AGGGGGGCGGCTGCGGGGAGAGACTGTGCTCAGATGATGATGATGATGATGATGATGATG 1801
DB 2267 GGGAGAGACCAAGCTGTGCTGCGGAGCTAGTGGCCATGCTGCAAGGAGATGACATCTGGCA 2326
QY 1802 TGAAGTACCTGAGAGAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1861
DB 2327 TGAATCACTGATATATCAATTAATATGTCACCGGAGACTGTGCTGCAAGAACTTTGG 2386
QY 1862 TGAAGTACCTGAGAGATGCTGGAAGATGATGATGATGATGATGATGATGATGATGATG 1921
DB 2387 TGAATCAAAACCTGTGCTGCAAGGTGTCTGATCTTGATCTTGATCTTGATCTTGATCT 2446
QY 1922 GGGTCTATGACAGCTTCAGGGGGGCTTCAGCAAGTCCCGTGAAGTGAACCGGACCTGAG 1981
DB 2447 TTGATGAGCAATACGAAACCCAGGGAGAAAGATCCCTATCCGTTGAGACAGCCCTGAG 2506
QY 1982 CCCTTAATACGAGCGGCTACTCTCCGAAAGCAAGTGTGAGCTTTGGCATCTTGCTCT 2041
DB 2507 CCATTGCCATGAGATCTTACCAACGACGAGGATGTGTGAGCTTTGGGATTTGATGT 2566
QY 2042 GGGAGACCTTCAGCGCTGGGGGCTCCCGCTATCCCACTTCAGCAATCAGACAGACGAG 2101
DB 2567 GGGAGGCTGCTGAGCTTTGGGGAGCAAGCTTATGAGGAGATGAGCAATCAGAGGTTATGA 2626
QY 2102 AGTTGTGAGAGAGGGGGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2161
DB 2627 AAGAGCATTTAGAGATGGGTATCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2686
QY 2162 GGCTCATGAGAGATGCTGGGCTTATGAGCTTGGGAGCGGCGCCAGCTTC 2211
DB 2687 AGCTCATGAGAGATGCTGGGCTTATGAGCTTGGGAGCGGCGCCAGCTTC 2736

```

Search completed: March 19, 2005, 19:55:55
Job time : 1181.22 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 03:43:36 ; Search time 369.746 Seconds
(without alignments)
9983.718 Million cell updates/sec

Title: US-10-660-763-1_COPY_72_2327

Perfect score: 2256
Sequence: 1 atggagctctctctctgagct.....gacccgaagcgagctcg 2256

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 1202784 segs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
6: /cgn2_6/prodata/1/ina/5B_COMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2256	100.0	2674	3 US-09-817-180-1	Sequence 1, Appl1
2	2256	100.0	2674	4 US-10-003-295-1	Sequence 1, Appl1
3	296.2	13.1	361	3 US-09-387-212-5	Sequence 5, Appl1
4	296.2	13.1	361	3 US-09-948-802-5	Sequence 5, Appl1
5	227.2	10.1	15297	3 US-09-817-180-3	Sequence 3, Appl1
6	227.2	10.1	15297	4 US-10-003-295-3	Sequence 3, Appl1
7	213	9.4	19152	4 US-09-949-016-12110	Sequence 12110, A
8	213	9.4	19152	4 US-09-949-016-15795	Sequence 15795, A
9	198	8.8	3370	4 US-09-814-815A-82	Sequence 82, Appl1
10	185.4	8.2	3921	4 US-09-949-016-628	Sequence 628, Appl1
11	185.4	8.2	3921	4 US-09-949-016-2120	Sequence 2120, Appl1
12	174.6	7.7	3386	4 US-09-949-016-5662	Sequence 5662, Appl1
13	174.6	7.7	3416	2 US-08-357-642A-2	Sequence 2, Appl1
14	174.6	7.7	3416	2 US-08-460-626-2	Sequence 2, Appl1
15	174.6	7.7	3416	4 US-09-016-434-1483	Sequence 1483, Appl1
16	174.6	7.7	4089	4 US-09-300-958A-13	Sequence 13, Appl1
17	170.4	7.6	1611	1 US-07-820-011A-3	Sequence 3, Appl1
18	170.4	7.6	1611	4 US-09-860-473-3	Sequence 3, Appl1
19	170.4	7.6	1611	4 US-09-444-711A-1	Sequence 1, Appl1
20	170.4	7.6	1611	4 US-09-444-711A-3	Sequence 3, Appl1
21	170.4	7.6	1611	5 PCT-US93-00445-3	Sequence 3, Appl1
22	170.4	7.6	2455	4 US-09-949-016-4411	Sequence 4411, Appl1
23	169.4	7.5	3623	1 US-08-306-691B-35	Sequence 35, Appl1
24	169.4	7.5	5434	4 US-09-949-016-927	Sequence 927, Appl1
25	169.4	7.5	5763	4 US-09-949-016-3759	Sequence 3759, Appl1
26	169	7.5	3018	4 US-09-949-016-1097	Sequence 1097, Appl1
27	169	7.5	3018	4 US-09-949-016-4749	Sequence 4749, Appl1

28	169	7.5	4871	4 US-09-799-451-448	Sequence 448, Appl1
29	168.2	7.5	2574	3 US-09-142-529-2	Sequence 2, Appl1
30	168.2	7.5	2574	4 US-10-045-428A-2	Sequence 2, Appl1
31	165.8	7.3	738	2 US-08-604-989A-8	Sequence 8, Appl1
32	165.8	7.3	1398	2 US-08-604-989A-9	Sequence 9, Appl1
33	165.8	7.3	1521	2 US-08-604-989A-10	Sequence 10, Appl1
34	165.8	7.3	1713	3 US-09-741-154-1	Sequence 1, Appl1
35	165.8	7.3	1942	2 US-08-604-989A-11	Sequence 11, Appl1
36	165.8	7.3	2000	3 US-08-426-509A-1	Sequence 1, Appl1
37	165.8	7.3	2000	4 US-08-232-545-1	Sequence 1, Appl1
38	163.6	7.3	2000	5 PCT-US95-05008-1	Sequence 1, Appl1
39	163.6	7.3	2449	4 US-09-949-016-1248	Sequence 1248, Appl1
40	163.6	7.3	5993	3 US-09-383-630-1	Sequence 1, Appl1
41	163.6	7.3	5993	3 US-09-383-630-2	Sequence 2, Appl1
42	162.6	7.2	1467	4 US-09-579-182-2	Sequence 2, Appl1
43	162.2	7.2	2440	1 US-08-160-861-2	Sequence 2, Appl1
44	162.2	7.2	2442	1 US-08-542-363-3	Sequence 3, Appl1
45	162.2	7.2	2442	3 US-09-100-089-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1					
US-09-817-180-1					
Sequence 1, Application US/09817180					
Patent No. 6340584					
GENERAL INFORMATION:					
APPLICANT: GAN, Weinlu et al.					
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC					
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES					
TITLE OF INVENTION: THREEOF					
FILE REFERENCE: CL001183					
CURRENT APPLICATION NUMBER: US/09/817,180					
CURRENT FILING DATE: 2001-03-27					
NUMBER OF SEQ ID NOS: 4					
SOFTWARE: FastSeq for Windows Version 4.0					
SEQ ID NO 1					
LENGTH: 2674					
TYPE: DNA					
ORGANISM: Human					
US-09-817-180-1					
Query Match					
Best Local Similarity 100.0%; Pred. No. 0;					
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	ATGGGCTTCCTCTGAGCGTGCAGCCCGGCGGCTCTGCAAAATGACAG	60		
DB	72	ATGGGCTTCCTCTGAGCGTGCAGCCCGGCGGCTCTGCAAAATGACAG	131		
QY	61	GAGGCGGCTTCCTCTGAGCGTGCAGCCCGGCGGCTCTGCAAAATGACAG	120		
DB	132	GAGGCGGCTTCCTCTGAGCGTGCAGCCCGGCGGCTCTGCAAAATGACAG	191		
QY	121	GACAGGAGTATGACAGCTGCTTCCACATGTCCTTGCAGACAGTGGGCGGCGG	180		
DB	192	GACAGGAGTATGACAGCTGCTTCCACATGTCCTTGCAGACAGTGGGCGGCGG	251		
QY	181	CGGCGGCTTCCTCTGAGCGTGCAGCCCGGCGGCTCTGCAAAATGACAG	240		
DB	252	CGGCGGCTTCCTCTGAGCGTGCAGCCCGGCGGCTCTGCAAAATGACAG	311		
QY	241	GAGGCGGCTTCCTCTGAGCGTGCAGCCCGGCGGCTCTGCAAAATGACAG	300		
DB	312	GAGGCGGCTTCCTCTGAGCGTGCAGCCCGGCGGCTCTGCAAAATGACAG	371		
QY	301	AACTGAGCTGCTCTATCCGGAACCGGCGGCTTCGCAAGCTTACGAGCGAGTGG	360		
DB	372	AACTGAGCTGCTCTATCCGGAACCGGCGGCTTCGCAAGCTTACGAGCGAGTGG	431		
QY	361	CAGCAGCTCAGAGGAGTCAACCAAGCCAGGCGGAGCACTTGAAGAGTGAAGAGC	420		

432 CAGCAGCTGACGAGAGCTCACCAAGACCCACAGCCAGGACATGTAGAAAGCTGAAGAC 491
421 CAGTACCGAGCTTGGGACCGGGACAGTGGCCCAAGCCAAAGGCAAGTACCAAGAGGGCAGC 480
492 CAGTACCGAGCTTGGGACCGGGACAGTGGCCCAAGCCAAAGGCAAGTACCAAGAGGGCAGC 551
481 AAAGACAAAGACCGGTACAAAGGCAAGGCAAGTATGTGGCGAGCTGTGGAAGCTCTTT 540
552 AAAGACAAAGACCGGTACAAAGGCAAGGCAAGTATGTGGCGAGCTGTGGAAGCTCTTT 611
541 GCTCACCAACCGGTATGTGTGGGCGTGGGCGTGGCGAGCTACCAACCAACCAAC 600
612 GCTCACCAACCGGTATGTGTGGGCGTGGGCGTGGCGAGCTACCAACCAACCAAC 671
601 CACCACTCTGTGCTGGGCGTGGGCGTGGGCGTGGGCGAGCTACCAACCAACCAAC 660
672 CACCACTCTGTGCTGGGCGTGGGCGTGGGCGTGGGCGAGCTACCAACCAACCAAC 731
661 TGCATCTGAAAGAGTCTGCAAGAAATACCTGGAAGTTCAGAGCTGTGCAAGTGAAG 720
732 TGCATCTGAAAGAGTCTGCAAGAAATACCTGGAAGTTCAGAGCTGTGCAAGTGAAG 791
721 GTGTGGCCATTCGACCGGAGATGGCTGACGTGGCGCATCCAGCTGAGGCTGAG 780
792 GTGTGGCCATTCGACCGGAGATGGCTGACGTGGCGCATCCAGCTGAGGCTGAG 851
781 TACCAAGGCTTCTGTGGGACAGTATGGGTCCGACCTGACCTCCACCTGTGTCAAGTTC 840
852 TACCAAGGCTTCTGTGGGACAGTATGGGTCCGACCTGACCTCCACCTGTGTCAAGTTC 911
841 GATGAGTCACTGCTTGAAGAGGAGTGAACCGCTGGAAGCTGGGAGCTCCAGCTGAAGAG 900
912 GATGAGTCACTGCTTGAAGAGGAGTGAACCGCTGGAAGCTGGGAGCTCCAGCTGAAGAG 971
901 CTGACTGTGAGAGCGCTGACGACACCGCTGACCTCAGTGAAGATGAGCTGTGGGCTC 960
972 CTGACTGTGAGAGCGCTGACGACACCGCTGACCTCAGTGAAGATGAGCTGTGGGCTC 1031
961 ACCGAATGTGTTCAGGCGGCGAGAGTGGTTAACGAGTGAACAGAGGCTCCGGAAT 1020
1032 ACCGAATGTGTTCAGGCGGCGAGAGTGGTTAACGAGTGAACAGAGGCTCCGGAAT 1091
1021 GAAGAGAGAGACACCAACCCCGGAGCGGGGTGCAAGTGTGGGCAAGAGCAAGTGTG 1080
1092 GAAGAGAGAGACACCAACCCCGGAGCGGGGTGCAAGTGTGGGCAAGAGCAAGTGTG 1151
1081 CAAGAGACACTGACGGGGCTGAGGTAGCGCTGTGACAGGCGCAAGCTGACGGGCCAG 1140
1152 CAAGAGACACTGACGGGGCTGAGGTAGCGCTGTGACAGGCGCAAGCTGACGGGCCAG 1211
1141 CAGGAGTGTGTGAGACCAAGCTGAGACCTGGGGCCCGGCGAGGCCCTGTGTGCTG 1200
1212 CAGGAGTGTGTGAGACCAAGCTGAGACCTGGGGCCCGGCGAGGCCCTGTGTGCTG 1271
1201 CTCTCAGAGATGACGCGCACTCAGCTCGTCTCCGAGAGGAGGAGGAGGAGG 1260
1272 CTCTCAGAGATGACGCGCACTCAGCTCGTCTCCGAGAGGAGGAGGAGGAGG 1331
1261 ACAACCAACGCTGAGATCTTAAAGGACACATCTCAGAAATCTTCCGCCCAAGTTCTG 1320
1332 ACAACCAACGCTGAGATCTTAAAGGACACATCTCAGAAATCTTCCGCCCAAGTTCTG 1391
1321 AACCTGTAACGATCTGAAAGGGGAAAGGTTCTTACGATCTTCTTGTCTCATCCACCTA 1380
1392 AACCTGTAACGATCTGAAAGGGGAAAGGTTCTTACGATCTTCTTGTCTCATCCCTA 1451
1381 CTGAGACCCAGAGCCCTCACCAAGAAAGTGTGTGTCTCTGCAAGGGCTGTGCC 1440
1452 CTGAGACCCAGAGCCCTCACCAAGAAAGTGTGTGTCTCTGCAAGGGCTGTGCC 1511
1441 AAGGACAGTGGGTCTGAAACATGAGGACCTGTGTGTGTGGGTGAGCAGATGAGCGGG 1500
1512 AAGGACAGTGGGTCTGAAACATGAGGACCTGTGTGTGTGGGTGAGCAGATGAGCGGG 1571

1501 AACTTTGGGAAAGTGTTCAGCGGAGCGCTGCGAGCGGACCAACACCCTGTGGCGGTGAAG 1560
1572 AACTTTGGGAAAGTGTTCAGCGGAGCGCTGCGAGCGGACCAACACCCTGTGGCGGTGAAG 1631
1561 TCTTGTGAGAGACGCTCCACCTGACCTCAAGGCGCAAGTTTCTACAGGAACGAGATC 1620
1632 TCTTGTGAGAGACGCTCCACCTGACCTCAAGGCGCAAGTTTCTACAGGAACGAGATC 1691
1621 CTGAAGCAGTACAGCCACCCCAACATCGTGCCTTCAATTGTGTCTGCAACCAAGAGCAG 1680
1692 CTGAAGCAGTACAGCCACCCCAACATCGTGCCTTCAATTGTGTCTGCAACCAAGAGCAG 1751
1681 CCCATCTACATGTCATGAGGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
1752 CCCATCTACATGTCATGAGGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1811
1741 GAGGGGCGCGCTGCGGAGTGAAGACTGTGTCAGAGATGTGGGAGATGACGTGTGCG 1800
1812 GAGGGGCGCGCTGCGGAGTGAAGACTGTGTCAGAGATGTGGGAGATGACGTGTGCG 1871
1801 ATGAGTACCTGAGAGCAAGTGTGATCCACCGGAGACTGGCTGTGGAATGCTG 1860
1872 ATGAGTACCTGAGAGCAAGTGTGATCCACCGGAGACTGGCTGTGGAATGCTG 1931
1861 GTGACAGAGAGATGTCTGAAAGATCAGTGACTTTGGGATGTCCGAGAGAGAGCGGAT 1920
1932 GTGACAGAGAGATGTCTGAAAGATCAGTGACTTTGGGATGTCCGAGAGAGAGCGGAT 1991
1921 GGGGTCTATGACGCTCAGGGGCGCTCAGACAAAGTCCCGTGAAGTGAACCGACCTGAG 1980
1992 GGGGTCTATGACGCTCAGGGGCGCTCAGACAAAGTCCCGTGAAGTGAACCGACCTGAG 2051
1981 GCCCTTAATACGCGCGCTACCTCCGAAACCGATCGTGTGAGCTTTGGCAATCTTGCTC 2040
2052 GCCCTTAATACGCGCGCTACCTCCGAAACCGATCGTGTGAGCTTTGGCAATCTTGCTC 2111
2041 TGGGAGACCTTACGCTGAGGAGGCGCTCCCTTATCCCAACCTGACAAATCAGCAGACCG 2100
2112 TGGGAGACCTTACGCTGAGGAGGCGCTCCCTTATCCCAACCTGACAAATCAGCAGACCG 2171
2101 GAGTTGTGAGAGAGGGGGCGCTGCTGCCCTGCCAGAGCTGTCTGATGCCGTGTC 2160
2172 GAGTTGTGAGAGAGGGGGCGCTGCTGCCCTGCCAGAGCTGTCTGATGCCGTGTC 2231
2161 AGGCTCATGAGACAGTGTGTGGCTCTTATGAGCTGTGGGCAAGCGGCTTTCAGCAATC 2220
2232 AGGCTCATGAGACAGTGTGTGGCTCTTATGAGCTGTGGGCAAGCGGCTTTCAGCAATC 2291
2221 TACCAAGAGCTGACAGCATCCGAAAGCGGCAATCGG 2256
2292 TACCAAGAGCTGACAGCATCCGAAAGCGGCAATCGG 2327

RESULT 2
US-10-003-295-1
; Sequence 1, Application US/10003295
; Patent No. 6686187
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001183DIV
; CURRENT APPLICATION NUMBER: US/10/003,295
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2674
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-003-295-1

applied

Query Match 100.0%; Score 2256; DB 4; Length 2674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCTTCTCTCTGAGCTGTGAGCCGCCAGGCGCAAGGGTCTGTGACGAATATGAC 60
DB 72 ATGGGCTTCTCTCTGAGCTGTGAGCCGCCAGGCGCAAGGGTCTGTGACGAATATGAC 131
QY 61 GAGGCGGAGCTTGTGTCTGAGGCGATGAGAAATGATGAGTGGCCAGCGGATCAAGAT 120
DB 132 GAGGCGGAGCTTGTGTCTGAGGCGATGAGAAATGATGAGTGGCCAGCGGATCAAGAT 191
QY 121 GACAGGAGATATGACAGACTGTCTTACACATGCTCTGTGAGACAGTGGGGCCAGAGC 180
DB 192 GACAGGAGATATGACAGACTGTCTTACACATGCTCTGTGAGACAGTGGGGCCAGAGC 251
QY 181 GGGGCGATCAGCCCTGACAGCCCATCAGTCACTGCTGGGCTGAGATCAACAGCAAACT 240
DB 252 GGGGCGATCAGCCCTGACAGCCCATCAGTCACTGCTGGGCTGAGATCAACAGCAAACT 311
QY 241 GAGGCGCTGAGCCGCTTGTCTGCGGACAGCAAGAGATCTGAATCAAGGCGCTGAGC 300
DB 312 GAGGCGCTGAGCCGCTTGTCTGCGGACAGCAAGAGATCTGAATCAAGGCGCTGAGC 371
QY 301 AAGCTGAGCTGCTCATCCGGGAAACGGACAGACTTGGCAAGCTTACAGCGAGCTG 360
DB 372 AAGCTGAGCTGCTCATCCGGGAAACGGACAGACTTGGCAAGCTTACAGCGAGCTG 431
QY 361 CAGCAGCTGACAGAGAGCTCACCAAGAACCCAGCAGGACCAATTGAGAACTGAAAGC 420
DB 432 CAGCAGCTGACAGAGAGCTCACCAAGAACCCAGCAGGACCAATTGAGAACTGAAAGC 491
QY 421 CAGTACCGAGCTTGGGACGGGACAGTCCCAAGCCAAAGCCAAAGTACAGAGGCGAGC 480
DB 492 CAGTACCGAGCTTGGGACGGGACAGTCCCAAGCCAAAGCCAAAGTACAGAGGCGAGC 551
QY 481 AAGGACAGAGACCTGTGACAAAGCCCAAGAGCAATATGTGCGACGCTGTGAAAGCTTT 540
DB 552 AAGGACAGAGACCTGTGACAAAGCCCAAGAGCAATATGTGCGACGCTGTGAAAGCTTT 611
QY 541 GCTCACCAACACGCTATGTGCTGGGCGTGGGGCTGCGGAGCTTACCAACAGCAACAC 600
DB 612 GCTCACCAACACGCTATGTGCTGGGCGTGGGGCTGCGGAGCTTACCAACAGCAACAC 671
QY 601 CACCAAGCTTGTGCTGCGGCTGCTGCGGCTGCTGAGGACCTGACAGAGAGATGCT 660
DB 672 CACCAAGCTTGTGCTGCGGCTGCTGCGGCTGCTGAGGACCTGACAGAGAGATGCT 731
QY 661 TGCATCTTGAAGAGATCTGTGACGAATATCTGAGATTAAGCAGCTGTGTGAGATGAG 720
DB 732 TGCATCTTGAAGAGATCTGTGACGAATATCTGAGATTAAGCAGCTGTGTGAGATGAG 791
QY 721 GTGGTGGCCATTACCGGGAGATGAGCTGTGAGCTGTGCGGCACTTGAAGGCTGAG 780
DB 792 GTGGTGGCCATTACCGGGAGATGAGCTGTGAGCTGTGCGGCACTTGAAGGCTGAG 851
QY 781 TACCAAGGCTTCTGTGCGACAGTATGAGTCCGACCTGACCTCCACCTGTGTCAAGTTC 840
DB 852 TACCAAGGCTTCTGTGCGACAGTATGAGTCCGACCTGACCTCCACCTGTGTCAAGTTC 911
QY 841 GATGAGTCACTGCTTGAAGAGGATGAAACGCTGAGAGCTTGGGAGCTCAAGTGAAGAG 900
DB 912 GATGAGTCACTGCTTGAAGAGGATGAAACGCTGAGAGCTTGGGAGCTCAAGTGAAGAG 971
QY 901 CTGACTGTGAGAGCGTGAAGCAACGCTGACCTCAAGTGAAGATGAGCTGTGAGGCGC 960
DB 972 CTGACTGTGAGAGCGTGAAGCAACGCTGACCTCAAGTGAAGATGAGCTGTGAGGCGC 1031
QY 961 ACCGAGATGTGTGTCAGGCGGACAGAGATGTTTACGAGCTGCAACAGAGCTCCGGAAT 1020
DB 1032 ACCGAGATGTGTGTCAGGCGGACAGAGATGTTTACGAGCTGCAACAGAGCTCCGGAAT 1091

QY 1021 GAAAGAGAAACACCCACCCCGGAGCGGGTGCAGCTGTGGGCAAGAGGCAAGTGTG 1080
DB 1092 GAAAGAGAAACACCCACCCCGGAGCGGGTGCAGCTGTGGGCAAGAGGCAAGTGTG 1151
QY 1081 CAAGAGCACTGACAGGGGCTGACAGTATGAGCTGTGACGACAGGCAAGCTGACAGGCC 1140
DB 1152 CAAGAGCACTGACAGGGGCTGACAGTATGAGCTGTGACGACAGGCAAGCTGACAGGCC 1211
QY 1141 CAGAGTGTCTGACAGCAAGCTGTGAGCAGCTGGGCTCCGCGAGAGCCCGCTGTGTCTG 1200
DB 1212 CAGAGTGTCTGACAGCAAGCTGTGAGCAGCTGGGCTCCGCGAGAGCCCGCTGTGTCTG 1271
QY 1201 CTCCTGCAAGATGACCGGCACTCAGCTGTCTGAGACAGAGAGGAGAGGGGGAAAG 1260
DB 1272 CTCCTGCAAGATGACCGGCACTCAGCTGTCTGAGACAGAGAGGAGAGGGGGAAAG 1331
QY 1261 ACACCAAGCTGAGATCTTAAAGCCACATCTCAGAACTCTCCGCCCAAGTCTG 1320
DB 1332 ACACCAAGCTGAGATCTTAAAGCCACATCTCAGAACTCTCCGCCCAAGTCTG 1391
QY 1321 AACCTGACCGATGGAAGGGGAAAGGCTTCTTGAAGATCTTCTGCTCATGACACTTA 1380
DB 1392 AACCTGACCGATGGAAGGGGAAAGGCTTCTTGAAGATCTTCTGCTCATGACACTTA 1451
QY 1381 CTGAGACCCAGAGGCGCTCCACCAAGAGAGTGTGTCTCTGACAGGGCTGTGCCC 1440
DB 1452 CTGAGACCCAGAGGCGCTCCACCAAGAGAGTGTGTCTCTGACAGGGCTGTGCCC 1511
QY 1441 AAGGACAAATGGGTGTGAACCAATGAGACCTGTGTTGGGTGAGAGCAATGAGCGGG 1500
DB 1512 AAGGACAAATGGGTGTGAACCAATGAGACCTGTGTTGGGTGAGAGCAATGAGCGGG 1571
QY 1501 AACTTTGGGAAAGTGTTCAGCGGACGCTTGGACCCGACCAACACTTGTGTGAGAG 1560
DB 1572 AACTTTGGGAAAGTGTTCAGCGGACGCTTGGACCCGACCAACACTTGTGTGAGAG 1631
QY 1561 TCTTGTGAGAGAGCGTCCACCTGACCTCAAGGGCAAGTGTCTTACAGAGGAGAGATC 1620
DB 1632 TCTTGTGAGAGAGCGTCCACCTGACCTCAAGGGCAAGTGTCTTACAGAGGAGAGATC 1691
QY 1621 CTGAAACAGTATCAGCAACCCCAACATCTGTGCTGTCAATGTTGTGTGCAACCCAGAGCAG 1680
DB 1692 CTGAAACAGTATCAGCAACCCCAACATCTGTGCTGTCAATGTTGTGTGCAACCCAGAGCAG 1751
QY 1681 CCCATCTAATCATGTATGAGACTTGTGACAGGGGGGAGACTTCTGACTTCTCCGACG 1740
DB 1752 CCCATCTAATCATGTATGAGACTTGTGACAGGGGGGAGACTTCTGACTTCTCCGACG 1811
QY 1741 GAGGGGGCGGCTGCGGGTGAAGACTTGTGCTGCAATGTGTGGGGATGCAAGCTGTGCGC 1800
DB 1812 GAGGGGGCGGCTGCGGGTGAAGACTTGTGCTGCAATGTGTGGGGATGCAAGCTGTGCGC 1871
QY 1801 ATGAGATCACTGAGAGCAAGTCTGATCAACCGGAGACTTGTGAGATGTCCGAGAGAGCCGAT 1860
DB 1872 ATGAGATCACTGAGAGCAAGTCTGATCAACCGGAGACTTGTGAGATGTCCGAGAGAGCCGAT 1931
QY 1861 GTGACAGAGAGAGATGTCTTGAAGATCACTTGTGAGATGTCCGAGAGAGCCGAT 1920
DB 1932 GTGACAGAGAGAGATGTCTTGAAGATCACTTGTGAGATGTCCGAGAGAGCCGAT 1991
QY 1921 GGGGTCTATGACAGCTTCAAGGGGCTTCAAGCAAGTCTCCGTGAAGTGAAGTGAAGCCGACCTGAG 2051
DB 1992 GGGGTCTATGACAGCTTCAAGGGGCTTCAAGCAAGTCTCCGTGAAGTGAAGTGAAGCCGACCTGAG 2051
QY 1981 GGCCTTAATCAACGCGGCTACTCTCCGAAAGGAGAGTGTGAGACTTGTGAGACTTGTGCTC 2040
DB 2052 GGCCTTAATCAACGCGGCTACTCTCCGAAAGGAGAGTGTGAGACTTGTGAGACTTGTGCTC 2111
QY 2041 TGGAGACCTTCAAGCTGTGGGGGCTTCCCTATCCCAACTCAGCAATCAGCAGACAG 2100
DB 2112 TGGAGACCTTCAAGCTGTGGGGGCTTCCCTATCCCAACTCAGCAATCAGCAGACAG 2171
QY 2101 GACTTGTGAGAGAGGGGGCGCTTGTGCTGTGCGCAGAGCTGTGTCTGTATGCGGTTC 2160

Position	Sequence	Position	Sequence
Db	2172 GAGTTTGTGGAGAAAGGGGGGGCCGTCCTGCGCCAGAGCTGTGTCCGATGCGCGTGTTC	2231	
Qy	2161 AGGCTCATGAGACCACTGCTGGGCGCTATAGAGCTGGGCAAGCGGCCCAAGTTTACAGACCATC	2220	
Db	2232 AGGCTCATGAGACCACTGCTGGGCGCTATAGAGCTGGGCAAGCGGCCCAAGTTTACAGACCATC	2231	
Qy	2221 TACCAAGAGCTGCAGAGCATCCGAAAGGGGCGATCGG	2256	
Db	2292 TACCAAGAGCTGCAGAGCATCCGAAAGGGGCGATCGG	2327	

```

RESULT 3
US-09-387-212-5
; Sequence 5, Application US/09387212A
; Patent No. 6309849
; GENERAL INFORMATION:
; APPLICANT: ROBINSON, KEITH E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; FILE REFERENCE: MNT-090
; CURRENT APPLICATION NUMBER: US/09/387,212A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-09-387-212-5

```

Query Match	13.1%	Score 296.2	DB 3	Length 361
Best Local Similarity	97.4%	Pred. NO. 6e-56		
Matches 333, Conservative	0	Mismatches	4	Indels 5, Gaps 3

QY	1613	CGAGATCTTGAAGTACAGACCCACCCCAACATGTGCGTCTCATTTGGTGTGCACCC	1672
Db	20	CNAGATCTTAAAGACATACAGCCACCCCAACATGTGCGTCTCATTTGGTGTGCACCC	79
QY	1673	AGAAGCAGCCCATCTACATTCGTCAATGGAAGCTTGTGCAAGGGGGGACACTTCCGACCTTCC	1732
Db	80	AGAAGCAGCCCATCTTACATTCGTCAATGGAAGCTTGTGCAAGGGGGGACACTTCTTACCTTCC	139
QY	1733	TCCGCACGAGAGGGGGCCCGCCTGCAGGTAAGACTCTGTGCAGATGTGTGGGGGATGCAG	1793
Db	140	TCCGCACGAGAGGGGGCCCGCCTGCAGGTAAGACTCTGTGCAGATGTGTGGGGGATGCAG	199
QY	1793	CTGCTGCATGAGATCACTTGAGAGCAAGTGTGCATCCACCGGACCTGGCTGTGCGA	1853
Db	200	CTGCTGCATGAGATCACTTGAGAGCAAGTGTGCATCCACCGGACCTGGCTGTGCGA	259
QY	1853	ACTGCTTGTGACAGAGAAATGTCTTGAAAG-ATCACTGACTTT-GGAGTGTCCCGAGA	1910
Db	260	ACTGCTTGTGACAGAGAAATGTCTTGAAAGATCACTGACTTTGGGAGATGTCCCGAGA	319
QY	1911	GGAGGCC---GATGGGTCCTAATGACGCTCAGGGGGGCTCGAG	1949
Db	320	GGAGGCCGATTTGGGGGTCTAATGACGCTCAGGGGGGCTTCGAG	361

RESULT 4
 US-09-948-802-5
 : Sequence 5, Application US/09948802
 : Patent No. 6465232
 : GENERAL INFORMATION:
 : APPLICANT: ROBINSON, KEITH E.
 : TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
 : TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
 : FILE REFERENCE: MN1-090
 : CURRENT APPLICATION NUMBER: US/09/948, 802
 : CURRENT FILING DATE: 2001-09-07

```

:
: PRIOR APPLICATION NUMBER: 09/387,212
:
: PRIOR FILING DATE: 1999-08-31
:
: NUMBER OF SEQ ID NOS: 18
:
: SOFTWARE: PatentIn Ver. 2.0
:
: SEQ ID NO 5
:
: LENGTH: 361
:
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-09-948-802-5

```

Query Match	13.1%	Score 296.2%	DB 3	Length 361
Best Local Similarity	97.4%	Pred. No. 6e-56		
Matches 333	Conservative 0	Mismatches 4	Indels 5	Gaps 3
QY	1613	CGAGAGTCCTGAAGCAGTACAGCCACCCCAACATCGTGCCTCTCATGTTGTTCTGCACCC	16772	
Db	20	CNAGAGTCCTGAAGCAGTACAGCCACCCCAACATCGTGCCTCTCATGTTGTTCTGCACCC	79	
QY	1673	AGAAAGCGCCCATCTTACATGTGTCAATGGAAGCTTGTGACGGGGGGGACCTTCTCTGACCTTCC	1732	
Db	80	AGAAAGCGCCCATCTTACATGTGTCAATGGAAGCTTGTGACGGGGGGGACCTTCTCTGACCTTCC	139	
QY	1733	TCCGACAGGAGGGGGGCCCCGCTGGGGGTGAAGACTCTGCTGCAGATGATGGGGGATGCAG	1792	
Db	140	TCCGACAGGAGGGGGGCCCCGCTGGGGGTGAAGACTCTGCTGCAGATGATGGGGGATGCAG	199	
QY	1793	CTGCTGGCATGGAATCCTTGAAGACAAAGTCTGTCATCCACCGGACCTTGGCTTCTCGGA	1852	
Db	200	CTGCTGGCATGGAATCCTTGAAGACAAAGTCTGTCATCCACCGGACCTTGGCTTCTCGGA	259	
QY	1853	ACTGCTCTGGTGAACAGAGAAAGAAATGTCTTGAAG-ATCAGTGACTTT-GGATGTCCGAGA	1910	
Db	260	ACTGCTCTGGTGAACAGAGAAAGAAATGTCTTGAAGATCAGTGAATCTTTGGGGAATGTCCGAGA	319	
QY	1911	GGAAAGCC---GATGGGGTCTATGACGCTCAGGGGGGCTCGAG	1949	
Db	320	GGAAAGCCGATGGGGGATCTATGACGCTCAGGGGGGCTCGAG	361	

```

RESULT 5
US-09-817-180-3
; Sequence 3, Application US/09817180
; Patent No. 6340584
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C001183
; CURRENT APPLICATION NUMBER: US/09/817,180
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Human
US-09-817-180-3

```

	Query Match	Best Local Similarity	Score	DB 3	Length
Matches	394	Conservative	70.1%	0	152
			10.1%	8	152
			Pred. 3.1e-40		
			Mismatches	8	Gaps
					1

US-09-949-016-628

Query Match 8.2%; Score 185.4; DB 4; Length 3921;
Best Local Similarity 55.1%; Pred. No. 3.3e-31;
Matches 385; Conservative 0; Mismatches 311; Indels 3; Gaps 1;

1546 CTGGTGGCCGTTGAAGCTTGTGAGAGAGCGCTCCCACTGACCTCAAGGCCAAGTTTCTA 1605
1546 CTGGTGGCCGTTGAAGCTTGTGAGAGAGCGCTCCCACTGACCTCAAGGCCAAGTTTCTA 1605
2037 CCGGTGGCCATCAAGAGCGCTGAAAGCCGGCTACACAGAGAGAGGAGGTGCTTCTC 2096
1606 CAGGAAGCAGAGATCTCTGAAGCAGTACAGCCACCCCAACATCGTGGCTCATTTGGTGC 1665
2097 GGGAGGCGCGGATCAATGGGCGGATTCAGCCACCAACATCATCCGCTAGAGGGCGTC 2156
1666 TGACCCAGAGAGCGCCATCTACATGTCATGAGAGCTTGTGAGAGGGGCGAATTCTCG 1725
2157 ATCTCCAAATACAGGCCATGATGATCATCATGATACATGAGATGAGAGGGCGCTGAG 2216
1726 ACCTTCCTCGGACGAGAGGGGGCGCGCTGGGGTGAAGCTGCTGACAGATGGTGGG 1785
2217 AAGTTCTTGGGAGAGAGATGAGGAGTTGAGCTGAGCTGAGTGGTGGCATGTGCGG 2276
1786 GATGAGCTGCTGGCAGTGAAGTACTGAGAGCAAGTGTGATCCACCGGAGCTGGCT 1845
2277 GGCATGAGCTGGCAGTGAAGTACTGAGAGCAAGTGTGATCCACCGGAGCTGGCT 2336
1846 GCTGGAATGCTGCTGGTGAAGAGATGTCCTGAAGTCAATGATCTTTGGGATGCTC 1905
2337 GCCCGCAACATCTCTCTCAACAGCAACCTGTGTCAGAGGTGTCGATCTTGGCTGTCC 2396
1906 CAGAGAGAGCGAGTGGGCTTATGCAAGCT---CAGGGGGCTCAGACAAAGTCCCGG 1962
2397 CGGTCTGAG 2456
1963 AAGTGAAGCAGCAGTGAAGGCTTAACTAGCGCGCTACTCTCCGAAAGCAGAGTGTG 2022
2457 CGGTGAG 2516
2023 AGCTTTGGCATCTTGTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2082
2517 AGCTTTGGCATCTTGTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2576
2083 AGCAATCAGACAGACAG 2142
2577 TCCAAACAG 2636
2143 TGTCTGATGCGGTGTTCAAGGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2202
2637 TGCCCTCGGCGCATCTACAGCTCATGATGATGATGATGATGATGATGATGATGATG 2696
2203 CCCAGCTTGAAGCAGATCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2241
2697 CCCAAGTTCGCTGACATCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2735

RESULT 11
US-09-949-016-2120
Sequence 2120, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2120
LENGTH: 3921
TYPE: DNA
ORGANISM: Human
US-09-949-016-2120

Query Match 8.2%; Score 185.4; DB 4; Length 3921;
Best Local Similarity 55.1%; Pred. No. 3.3e-31;
Matches 385; Conservative 0; Mismatches 311; Indels 3; Gaps 1;

1546 CTGGTGGCCGTTGAAGCTTGTGAGAGAGCGCTCCCACTGACCTCAAGGCCAAGTTTCTA 1605
1546 CTGGTGGCCGTTGAAGCTTGTGAGAGAGCGCTCCCACTGACCTCAAGGCCAAGTTTCTA 1605
2037 CCGGTGGCCATCAAGAGCGCTGAAAGCCGGCTACACAGAGAGAGGAGGTGCTTCTC 2096
1606 CAGGAAGCAGAGATCTCTGAAGCAGTACAGCCACCCCAACATCGTGGCTCATTTGGTGC 1665
2097 GGGAGGCGCGGATCAATGGGCGGATTCAGCCACCAACATCATCCGCTAGAGGGCGTC 2156
1666 TGACCCAGAGAGCGCCATCTACATGTCATGAGAGCTTGTGAGAGGGGCGAATTCTCG 1725
2157 ATCTCCAAATACAGGCCATGATGATCATCATGATACATGAGATGAGAGGGCGCTGAG 2216
1726 ACCTTCCTCGGACGAGAGGGGGCGCGCTGGGGTGAAGCTGCTGACAGATGGTGGG 1785
2217 AAGTTCTTGGGAGAGAGATGAGGAGTTGAGCTGAGCTGAGTGGTGGCATGTGCGG 2276
1786 GATGAGCTGCTGGCAGTGAAGTACTGAGAGCAAGTGTGATCCACCGGAGCTGGCT 1845
2277 GGCATGAGCTGGCAGTGAAGTACTGAGAGCAAGTGTGATCCACCGGAGCTGGCT 2336
1846 GCTGGAATGCTGCTGGTGAAGAGATGTCCTGAAGTCAATGATCTTTGGGATGCTC 1905
2337 GCCCGCAACATCTCTCTCAACAGCAACCTGTGTCAGAGGTGTCGATCTTGGCTGTCC 2396
1906 CAGAGAGAGCGAGTGGGCTTATGCAAGCT---CAGGGGGCTCAGACAAAGTCCCGG 1962
2397 CGGTCTGAG 2456
1963 AAGTGAAGCAGCAGTGAAGGCTTAACTAGCGCGCTACTCTCCGAAAGCAGAGTGTG 2022
2457 CGGTGAG 2516
2023 AGCTTTGGCATCTTGTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2082
2517 AGCTTTGGCATCTTGTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2576
2083 AGCAATCAGACAGACAG 2142
2577 TCCAAACAG 2636
2143 TGTCTGATGCGGTGTTCAAGGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2202
2637 TGCCCTCGGCGCATCTACAGCTCATGATGATGATGATGATGATGATGATGATGATG 2696
2203 CCCAGCTTGAAGCAGATCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2241
2697 CCCAAGTTCGCTGACATCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2735

RESULT 12
US-09-949-016-5662
Sequence 5662, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 5662
LENGTH: 3386
TYPE: DNA
ORGANISM: Human
US-09-949-016-5662

Query Match 7.7%; Score 174.6; DB 4; Length 3386;
Best Local Similarity 54.7%; Pred. No. 7.5e-29;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;

QY 1498 GGGAACTTTGGCGAAGTTCAGCGGACGCTCGAGCCGACAAACCTCTGGGGCGTG 1557
DB 1386 GGGAGAGTCTATGAAAGGTGTCTACAAATCACAAGGGAGAAATCAATGTAGCTGTC 1445
QY 1558 AAGCTTGTGAGAGAGCGCTCCACCTGACCTCAAGGCCAAGTTTCTACAGAAAGGAG 1617
DB 1446 AAGAGCTGCAAGAAAGCTGCACTGAGCAACAGAGAAAGTTTATGACGAGGAGTG 1505
QY 1618 ATCTGAGAGAGTACAGCCCAACCCCAATCATGCTGCTCATTTGGTGTCTGACCCAGAG 1677
DB 1506 ATCATGAAAGAACTCGACCAACCCGACATGCTGAAAGCTGATCGGCAT---CATGAAAG 1562
QY 1678 CAGCCATCTACATGCTGATGAGCTTGTGAGGGGGGCACTTCTGACCTTCTCCG 1737
DB 1563 GAGCCCACTGAGATCATGAGAAATGTATCCATATGGGAGCTGGGCACTACCTGAG 1622
QY 1738 ACGAGGGGGGGCGGCGGCGGAGAAAGTCTGCTGAGATGAGGGGAGTGAAGCTGCT 1797
DB 1623 CGGAACAAGAACTCTCGAAGAGTGTCAACCTGTGTGATCTGCACTGCAAGATAGCAA 1682
QY 1798 GGCATGAGTACCTGAGAGCAAGTGTGATCCAGCGGAGCTGGCTCGGAACTGC 1857
DB 1683 GGCATGAGTACCTGAGAGCAAGTGTGATCCAGCGGAGCTGGCTCGGAACTGC 1742
QY 1858 CTGGTACAGAGAAAGTGTCTGAGAGTCACTGATCTTTGGGATGTCCGAGAGAAAGCC 1917
DB 1743 CTGGTACCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1802
QY 1918 GATGGGCTATGAGAGCTGAGGGGGGCTCAGCAAGTCCCGGAGAGTGAAGCAAGCT 1977
DB 1803 GAGAGAGTAT---TACAAAGCTCTGTGATGTGTCTCCCATCAATGATGTCCCA 1859
QY 1978 GAGGCTTATACAGGCGCTACTCTCCGAAAGCAGAGTGTGAGAGTGTGAGAGTGTG 2037
DB 1860 GAGTCCATTAAGTCTCGAGGCTTCAAGCAGCAGTGAAGTGTGAGAGTGTGAGAGTGTG 1919
QY 2038 CTCTGGAGAGCTTCAAGCTTGGGGGCTCCCTCTATCCCACTCAGCAATCAGCAGACA 2097
DB 1920 ATGTGGAGAGTCTGAGCTTGGGAGCAGCCCTTCTTCTGCTGAGAGAAAGAGAGTGC 1979
QY 2098 CGGAGAGTGTGAGAGAGGGGGGCTGTGCGCTGAGAGAGTGTGAGAGTGTGAGAGTGTG 2157
DB 1980 ATCGGAGTGTGAG 2039
QY 2158 TTACAGCTATGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 2214
DB 2040 TATACCTCATGAG 2096

RESULT 13
US-08-357-642A-2
Sequence 2, Application US/08357642A
PATENT NO. 5837524
GENERAL INFORMATION:
APPLICANT: Sima Lev
APPLICANT: Joseph Schlessinger
TITLE OF INVENTION: PYK2 RELATED PRODUCTS

TITLE OF INVENTION: AND METHODS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSER: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/357,642A
FILING DATE: December 15, 1994
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 209/070
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3416
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic
US-08-357-642A-2

Query Match 7.7%; Score 174.6; DB 2; Length 3416;
Best Local Similarity 54.7%; Pred. No. 7.5e-29;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;
QY 1498 GGGAACTTTGGCGAAGTTCAGCGGACGCTCGAGCCGACAAACCTCTGGGGCGTG 1557
DB 1416 GGGAGAGTCTATGAAAGGTGTCTACAAATCACAAGGGAGAAATCAATGTAGCTGTC 1475
QY 1558 AAGCTTGTGAGAGAGCGCTCCACCTGACCTCAAGGCCAAGTTTCTACAGAAAGGAG 1617
DB 1476 AAGAGCTGCAAGAAAGCTGCACTGAGCAACCAAGAGAAAGTTTATGAGCAGAGAGTG 1535
QY 1618 ATCTGAGAGAGTACAGCCCAACCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 1677
DB 1536 ATCATGAAAGAACTTCACCAACCCGACATGTAAGTGTGAGAGTGTGAGAGTGTGAGAG 1592
QY 1678 CAGCCATCTACATGCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1737
DB 1593 GAGCCCACTGATCATCATGAAATGTATCCATATGGGAGAGTGTGAGAGTGTGAGAG 1652
QY 1738 ACGAGAGGGGGCGGCGGCGGAGAAAGTGTGCTGCAAGTGTGAGAGTGTGAGAGTGTGAG 1797
DB 1653 CGGAACAAGAAAGTCTCGAAGAGTGTCAACCTGTGTGATCTGCACTGAGATATCAAA 1712
QY 1798 GGCATGAGTACCTGAGAGAGCAAGTGTGATCCAGCGGAGCTGTGCTCGGAAGTGC 1857
DB 1713 GGCATGAGTACCTGAGAGAGCAAGTGTGATCCAGCGGAGCTGTGCTCGGAAGTGC 1772
QY 1858 CTGGTACAGAGAAAGTGTCTGAGAGTCAAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1917
DB 1773 CTGGTACCTCCCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1832
QY 1918 GATGGGCTATGAGAGCTGAGGGGGCTCAGCAAGTCCCGTGAAGTGAAGCAGCACT 1977

Db 1833 GAGGAGACTAT--TACAAAGCCTCTGTGATCTGCTCTCCCAATCAATGATGTCCTCA 1889
Qy 1978 GAGGCTTAACTACGAGCGCTACTCTCCGAAGAAGCATGTGAGCTTTGGCATCTTG 2037
Db 1890 GAGTCCATTAACTTCCGAGCGCTTACAGACAGCGAGTGAAGCTTGGAATGTTCCGCGGTGC 1949
Qy 2038 CTGTGGAGACCTTCAGCTGCGGGGCTCCCTTATCCCACTCAGCAATCAGCAGACA 2097
Db 1950 ATGTGGAGATCTCTGAGCTTTGGGAAGCAGCCCTTCTTGCTGAGAACAAAGATGTC 2009
Qy 2098 CGGAGATTTGTGAGAGAGGGGGCGCTGTCGCTGCGCAGAGCTGTGTCGTATGCGGTG 2157
Db 2010 ATCGGGGTCTGTGAGAAAGAGACCGGCTGCCAAGCTGATCTGTTCACCGGTCTT 2069
Qy 2158 TTCAGGCTCATGAGAGAGTGTGGGCTTATGAGCTTGGGAGCGGCCCAAGCTTCAGC 2214
Db 2070 TATACCTCATGACCGGCTGTGGAGCTAGACCCGAGTGAACCGGCCCGCTTCACG 2126

RESULT 14
US-08-460-626-2

/ Sequence 2, Application US/08460626
/ Patent No. 5837815
/ GENERAL INFORMATION:
/ APPLICANT: SIMA LEV
/ APPLICANT: JOSEPH SCHLESSINGER
/ TITLE OF INVENTION: PYK-2 RELATED PRODUCTS AND
/ TITLE OF INVENTION: METHODS
/ NUMBER OF SEQUENCES: 25
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Lyon & Lyon
/ STREET: 633 West Fifth Street
/ CITY: Los Angeles
/ STATE: California
/ COUNTRY: U.S.A.
/ ZIP: 90071-2066
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
/ MEDIUM TYPE: Storage
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: IBM P.C. DOS 5.0
/ SOFTWARE: Word Perfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/460,626
/ FILING DATE: June 2, 1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/357,642
/ FILING DATE: December 15, 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warburg, Richard J.
/ REGISTRATION NUMBER: 32,327
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (213) 489-1600
/ TELEFAX: (213) 955-0440
/ TELEX: 673510
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3416
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ US-08-460-626-2

Query Match 7.7%; Score 174.6; DB 2; Length 3416;
Best Local Similarity 54.7%; Pred. No. 7.5e-29;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;

Qy 1498 GGGAACTTGGGGAAGTGTTCAGCGGCGCTGCGAGCGAGCAACCCCTGTGGGGGTG 1557

Db 1416 GGGAGGTCTATGAAGGTGTCTACACAATCAAAAGGGGAGAAATCAATGATGCTGC 1475
Qy 1558 AAGTCTTGTGAGAGAGCGTCCCACTGACTTCAAGGCCAAGTTCTTACAGAGAGAG 1617
Db 1476 AAGACCTGCAAGAAAGACTGCACTCTGACAAACAGAGAAAGTTCAATGAGCGAGGAGTG 1535
Qy 1618 ATCTGGAACAGTACAGCAACCCCAACATGTGCGCTCATTTGATGTCTGACCCCAAG 1677
Db 1536 ATCATGAAAGACCTCAACCAACCGACATGTGAGAGCTATGAGGAT--CATTGAAGAG 1592
Qy 1678 CAGCCCATTAACATGTCATGAGACTTGTGACAGGGGGGAGACTTCTGACTCTCCG 1737
Db 1593 GAGCCCACTGATATCATGGAATTGTATCCCTATGGGAGAGTGGCCACTACCTGGAG 1652
Qy 1738 ACGAGAGGGGCGGCTGCGGGTGAAGACTTGTCTGACAGATGATGGGGAGATGCACTGCT 1797
Db 1653 CGGAAACAGAACTCCCTGAAGGTGCTACCCCTGTGCTACTACATGACATATACAA 1712
Qy 1798 GGCATGGAATACCTGAGAGCAAGTGTGCTATCCACCGGGACCTGCTGAGAACTGC 1857
Db 1713 GGCATGCTTACCTGAGAGCATCAACTGCTGACAGAGGACATTTCTGCGAAATGTC 1772
Qy 1858 CTGTGACAGAGAGAAATGCTCTGAAGATCAGTGAATTTGGAGATGTCGAGAGAGAGCC 1917
Db 1773 CTGTGCTCTCCCTGATGTGTGTAAGCTGGGGGACTTGTGCTTCCGCTACATTGAG 1832
Qy 1918 GATGGGTCTATGACAGCTTCAGGGGGCTTCAGACAAAGTCCCGTGAATGAGACCGACT 1977
Db 1833 GACGAGAGACTAT--TACAAAGCTGTGACTGCTGCCATCAATGATGTCCTCCA 1889
Qy 1978 GAGGCTTAACTACGCGCGCTACTCTCCGAAGAAGCATGTGAGACTTTGSCATCTTG 2037
Db 1890 GAGTCCATTAACTTCCAGCGCTTCAAGACAGACAGTGAAGTGTGCGCGGTGC 1949
Qy 2038 CTGTGGAGACCTTCAGCTGCGGGGCTCCCTTATCCCACTCAGCATCAGACAGACA 2097
Db 1950 ATGTGGAGATCTGAGCTTTGGGAAGAGCCCTTCTTGTGCTGAGAAAGATGTC 2009
Qy 2098 CGGAGATTTGTGAGAGAGGGGGCGCTGTGCTGCCAGAGCTGTCTGATGCGGTG 2157
Db 2010 ATCGGGGTCTGTGAGAAAGAGACCGGCTGCCAAGCTGATCTGTCCACCGGTCTT 2069
Qy 2158 TTCAGGCTCATGAGAGAGTGTGGGCTTATGAGCTTGGGAGCGGCCCAAGCTTCAGC 2214
Db 2070 TATACCTCATGACCGGCTGTGGAGCTAGACCCGAGTGAACCGGCCCGCTTCACG 2126

RESULT 15
US-09-016-434-1483

/ Sequence 1483, Application US/09016434
/ Patent No. 6500938
/ GENERAL INFORMATION:
/ APPLICANT: Janice Au-Young
/ APPLICANT: Jeffrey J. Seilhamer
/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
/ TITLE OF INVENTION: PATHWAY GENE EXPRESSION
/ NUMBER OF SEQUENCES: 1490
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 PORTER DRIVE
/ CITY: PALO ALTO
/ STATE: CALIFORNIA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/016,434
/ FILING DATE: HEREWITH
/ CLASSIFICATION:

Search completed: March 19, 2005, 14:24:17
Job time : 373.746 secs

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1483:
SEQUENCE CHARACTERISTICS:
LENGTH: 3416 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g988304
US-09-016-434-1483

Query Match 7.7%; Score 174.6; DB 4; Length 3416;
Best Local Similarity 54.7%; Pred. No. 7.5e-29;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;

QY 1498 GGGAACTTGGCGAAGTGTTCAGCGGACGCGCTGCGAGCAACACCTGTGGCGGTG 1557
DB 1416 GGGAGAGTCTATGAGGTGTCTACACAATACAAAGGGAGAAATCATATGCTGTC 1475
QY 1558 AAGCTTGTGAGAGAGAGCTGCCACCTGACCTCAAGGCCAAGTTTCTACAGAAAGCAGG 1617
DB 1476 AAGACCTGCAAGAAAGACTGACTCTGGAACAAGAGAAAGTTTCATGACGAGGCGAGTG 1535
QY 1618 ATCTGAGAGAGTACAGACCCCAACATGTGTGCTCATTTGTGTCTGACCCAGAG 1677
DB 1536 ATCATGAAAGACCTGACACCCGACATGATGAGTATCGGCAT--CATTTAAGAG 1592
QY 1678 CAGGCCATCTACATCGTCATGAGCTGTGAGAGGGGGCGACTTCTGACCTTCTCCGCG 1737
DB 1593 GAGCCCACTGATCATATGAAATTTATTCCTATGGAGAGCTGGGCCACTACTGAGAG 1652
QY 1738 ACGGAGGGGGCCCGCTCGGGGTGAAGACTCTGCTGCAAGTGTGGGGAGTGCAGCTGCT 1797
DB 1653 CGGAACAAGAACTCCCTGAAGGTGCTCACCCCTGCTGTACTGACATGACAGATATGCA 1712
QY 1798 GGCATGAGATACCTGAGAGAGCAAGTGTGATCCACCGGAGCTGTGCTCGGAATGCG 1857
DB 1713 GCCATGGCTTACCTGAGAGACATCACTGCTGCAAGAGACATGTCTGTCCGAAATC 1772
QY 1858 CTGGTGACAGAGAAATGTCTCTGAAGATCAGTACTTTGGGATGTCCCGAGAGAGAGCC 1917
DB 1773 CTGGTGGCTTCCCTGAGTGTGTAAGCTGGGGAGCTTTGGTCTTTCCGGTACATTGAG 1832
QY 1918 GATGGGCTTATGAGAGCTCAAGGGGAGCTCAGACAACTGCCGTGAAGTGAACGCACT 1977
DB 1833 GACGAGAGACTAT--TACAAAGCTCTGTGACTGTCTCTCCCAATCAATGATGTCCCA 1889
QY 1978 GAGGCCCTTAACTAGCGGCCGCTACTCTCCGAAAGCAGCTGTGAGAGCTTTGGCATTTG 2037
DB 1890 GAGTCCATTAACTTCCAGCGTTTCCAGACAGCCAGTACGTCGTGATGTTGCGCGTGTGC 1949
QY 2038 CTCTGGGAGACCTTCAAGCTGTGGGGGCTCCCTCTATCCCACTCAGCAATCAGACAGCA 2097
DB 1950 ATGTGGAGATCTGAGCTTTGGGAAGACGCTTTCTTCTGGCTGAGAAACAAGATGTC 2009
QY 2098 CGGAGATTGTGAGAAAGGGGGGCGTCTGTGCTGCCAGAGCTGTCTGTATGCCGTG 2157
DB 2010 ATCGGGGTGCTGAGAGAAAGAGACCGGCTGCCAAGCTGATCTGTGCCAGCGGTCTT 2069
QY 2158 TTCAGGCTCATGAGCAGTGTGGGCTTATGAGCTTGGGACCGGCCCAAGCTTCAAGC 2214
DB 2070 TATACCTCATGACCGCTGCTGGAACTACGACCCCAAGTACCGGCGCCGCTTCACC 2126

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 03:22:22 ; Search time 6515.86 Seconds

(without alignment)
13179.075 Million cell updates/sec

Title: US-10-660-763-1_COPY_72_2327

Perfect score: 2256

Sequence: 1 atgggctctctctctgcgact.....gcattccgaagcgacgcg 2256

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

EST: *
1: gb_esc1: *
2: gb_esc2: *
3: gb_esc3: *
4: gb_esc4: *
5: gb_esc5: *
6: gb_esc6: *
7: gb_esc7: *
8: gb_esc8: *
9: gb_esc9: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1732.8	76.8	2461	3	CR624741 full-length
2	1002	44.4	1189	3	CR593957 full-length
3	923.6	40.9	936	5	BX356089 BX356089
4	885.4	39.2	3373	3	AK004587 Mus muscu
5	848.8	37.6	936	1	AL542976
6	805.8	35.7	1029	4	BM456755
7	783.6	34.7	968	5	BX359010
8	773.4	34.3	939	5	BX381726
9	771.6	34.2	948	5	BQ708270
10	751.4	33.3	980	5	BX356088
11	716.8	31.8	922	5	BUI56855
12	710.6	31.5	1114	4	BM543870
13	703.2	31.2	847	4	BUI84736
14	697.6	30.9	989	5	BX359009
15	669.4	29.7	973	5	BX327963
16	662.4	29.4	913	3	AL569718
17	642	28.5	863	3	CR622908
18	640.4	28.4	861	1	AL540776
19	635.4	28.2	802	1	BP157303
20	630	27.9	848	1	AL540775
21	629	27.9	629	7	CV030188
22	601.6	26.7	882	6	BY704806
23	600.8	26.6	652	7	CN298839
24	599.8	26.6	880	5	BX381725

25	599.4	26.6	962	5	BO679512	BO679512
26	597	26.5	938	5	BO712558	BO712558
27	587.6	26.0	911	5	BU191031	BU191031
28	581.6	25.8	812	7	CK788677	CK788677
29	575.8	25.5	782	7	CK357411	CK357411
30	574.8	25.5	912	5	BO924403	BO924403
31	574.4	25.5	638	5	BO807439	BO807439
32	570.8	25.3	1077	5	BM922603	BM922603
33	570.2	25.3	787	6	CB990282	CB990282
34	551.2	24.4	662	2	AW701492	AW701492
35	547.2	24.3	674	7	CK943999	CK943999
36	528.4	23.4	580	2	BP332887	BP332887
37	526.2	23.3	943	2	BF159511	BF159511
38	517	22.9	805	5	BX544743	BX544743
39	515.8	22.9	675	7	CF731836	CF731836
40	506.4	22.4	573	2	AW658298	AW658298
41	498	22.1	582	5	BP256574	BP256574
42	497.6	22.1	643	2	BE917487	BE917487
43	496.6	22.0	748	4	BI910442	BI910442
44	494	21.9	732	5	BQ015500	BQ015500
45	475.6	21.1	600	4	BG803019	BG803019

ALIGNMENTS

RESULT 1	CR624741	2461 bp	mRNA	linear	HTC 21-JUL-2004
LOCUS	full-length cDNA clone CS01069YK21 of Placenta Cot 25-normalized				
DEFINITION	of Homo sapiens (human).				
ACCESSION	CR624741				
VERSION	CR624741.1	GI:50505548			
KEYWORDS	HTC; CNSLT; cDNA.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
REMARK	Contact : Feng Liang Email : fliang@life.rockefeller.edu http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue				
REFERENCE	Genoscope. 2 (bases 1 to 2461)				
AUTHORS	Direct Submission				
TITLE	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage ;				
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
COMMENT	- Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.				
FEATURES	Location/Qualifiers				
source	1..2461				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CS01069YK21"				
	/issue_type="Placenta Cot 25-normalized"				
	/plasmid="pCMVSPORT_6"				
ORIGIN					
Query Match	76.8% ; Score 1732.8 ; DB 3 ; Length 2461 ;				
Best Local Similarity	84.9% ; Pred. No. 0 ;				
Matches 2131 ; Conservative	0 ; Mismatches 2 ; Indels 377 ; Gaps 2 ;				
Qy	1 ATGGGCTTCTCTTCTGAGCTGTCAGCCCCCAAGGCGCAGGGGCTCTGAGCAAAATGACAG 60				
Db	32 ATGGGCTTCTCTTCTGAGCTGTCAGCCCCCAAGGCGCAGGGGCTCTGAGCAAAATGACAG 91				

QY 61 GAGGCGGAGCTTGTCTACTGAGGGCATGAGAAAGTGAAGTGGCCGAGCGGCTCAAGAT 120
DB 92 GAGGCGGAGCTTGTCTACTGAGGGCATGAGAAAGTGAAGTGGCCGAGCGGCTCAAGAT 151
QY 121 GACAGGAGATGATGACAGGACTGCTTCAACAATGTCCCTGAGGACAGTGGGGCCAGAC 180
DB 152 GACAGGAGATGATGACAGGACTGCTTCAACAATGTCCCTGAGGACAGTGGGGCCAGAC 211
QY 181 CGGGCCATCAAGCCCTTGAAGCCCATGATCAATCTTGGGCTGAATCAACAGCCAAAT 240
DB 212 CGGGCCATCAAGCCCTTGAAGCCCATGATCAATCTTGGGCTGAATCAACAGCCAAAT 271
QY 241 GAGGCGCTGAGCGGCTTGTCTGAGGAGCAAGGAGGATGAACTCAAGGCGCCCTGAGC 300
DB 272 GAGGCGCTGAGCGGCTTGTCTGAGGAGCAAGGAGGATGAACTCAAGGCGCCCTGAGC 331
QY 301 AAGCTGAGCTGTCTATCCGGGAAACGGCAGAGCTTTCGCAAGAATTACAGCGAGCATGG 360
DB 332 AAGCTGAGCTGTCTATCCGGGAAACGGCAGAGCTTTCGCAAGAATTACAGCGAGCATGG 391
QY 361 CAGCACTGACAGAGAGCTCAACCAAGCCCAAGCCAGGACATTTGAAGAGCTGAAGAGC 420
DB 392 CAGCACTGACAGAGAGCTCAACCAAGCCCAAGCCAGGACATTTGAAGAGCTGAAGAGC 451
QY 421 CAGTACCGAGCTGTGCAACGGGACAGTGCACCAAGCCAAAGGCAAGTACAGAGAGCAGC 480
DB 452 CAGTACCGAGCTGTGCAACGGGACAGTGCACCAAGCCAAAGGCAAGTACAGAGAGCAGC 511
QY 481 AAGAGCAAGGACCGTGAACAAGGCCAAGACCAATATGTGCGAGCCTGTGAAGCTCTTT 540
DB 512 AAGAGCAAGGACCGTGAACAAGGCCAAGACCAATATGTGCGAGCCTGTGAAGCTCTTT 571
QY 541 GCTCAACAACAACCGCTATGTGCTGGGCGTGGGGCTGGCGAGTAAACCAACGAGCAAC 600
DB 572 GCTCAACAACAACCGCTATGTGCTGGGCGTGGGGCTGGCGAGTAAACCAACGAGCAAC 631
QY 601 CACCACTCTGCTGCGCGGCTGTGCGGCTCACTGCAAGGACTGCAAGAGAGATGGCT 660
DB 632 CACCACTCTGCTGCGCGGCTGTGCGGCTCACTGCAAGGACTGCAAGAGAGATGGCT 691
QY 661 TGCATCTTGAAGAGATCTCTGAGAAATCTGAGAAATTAAGCAAGCTGTGAGATGAG 720
DB 692 TGCATCTTGAAGAGATCTCTGAGAAATCTGAGAAATTAAGCAAGCTGTGAGATGAG 751
QY 721 GTGGTGGCCATTTACCGGGAGATGGCTGCAAGCTGTGCCGCAATCCAGCCTGAGGCTGAG 780
DB 752 GTGGTGGCCATTTACCGGGAGATGGCTGCAAGCTGTGCCGCAATCCAGCCTGAGGCTGAG 811
QY 781 TACCAAGGCTTCTGCGACAGTATGGGTCGCACTGACGTCCACCTGTGTCAAGTTC 840
DB 812 TACCAAGGCTTCTGCGACAGTATGGGTCGCACTGACGTCCACCTGTGTCAAGTTC 871
QY 841 GATGAGTCACTGCTTGAAGAGAGGTGAACCGCTGAGGCTTGGGAGCTTCAAGCTGAACGAG 900
DB 872 GATGAGTCACTGCTTGAAGAGAGGTGAACCGCTGAGGCTTGGGAGCTTCAAGCTGAACGAG 931
QY 901 CTGACCTGTGAAGAGCGTGAACAAGCTGACCTCAATGAACAATGAGCTGTGGGCT 960
DB 932 CTGACCTGTGAAGAGCGTGAACAAGCTGACCTCAATGAACAATGAGCTGTGGGCT 991
QY 961 ACCGAGATGGTGTTCAGAGCGGAGAGATGTTACGAGAGCTGCAACAGAGCTCCGGAAT 1020
DB 992 ACCGAGATGGTGTTCAGAGCGGAGAGATGTTACGAGAGCTGCAACAGAGCTCCGGAAT 1051
QY 1021 GAAAGAGAGAACCCCAACCCCGGAGACGGGCTCAAGCTGTGGGCAAGAGGCAAGTCTG 1080
DB 1052 GAAAGAGAGAACCCCAACCCCGGAGACGGGCTCAAGCTGTGGGCAAGAGGCAAGTCTG 1111
QY 1081 CAAAGAGCACTGAGAGGGGCTGCAAGTATGCGCTGTGCAAGGCAAGCTGCAAGGCTCCAG 1140
DB 1112 CAAAGAGCACTGAGAGGGGCTGCAAGTATGCGCTGTGCAAGGCAAGGCTGCAAGGCTCCAG 1171

QY 1141 CAGAGTGTCTGACAGCAAGCTGAGACCTTGGGCGCCGAGGAGCCCGCTGTGCTG 1200
DB 1172 CAGAGTGTCTGACAGCAAGCTGAGACCTTGGGCGCCGAGGAGCCCGCTGTGCTG 1231
QY 1201 CTCTGCAAGATGACCGGCACTTCCACCTGTGTCTGAGACAGAGATGAGAGGGGGAGAG 1260
DB 1232 CTCTGCAAGATGACCGGCACTTCCACCTGTGTCTGAGACAGAGATGAGAGGGGGAGAG 1291
QY 1261 ACACCAAGCTGAGATCTTTAAGAGCCACATCTGAGAAATCTTCCGCCCCAAGTTCTC- 1319
DB 1292 ACACCAAGCTGAGATCTTTAAGAGCCACATCTGAGAAATCTTCCGCCCCAAGTTCTG 1351
QY 1320 ----- 1319
DB 1352 CTCTCCACCGGTGAGCTATTCGAGAGGTGAGAAAGCCCTGTGATGAGAGCTGTG 1411
QY 1320 ----- 1319
DB 1412 TACCAAGGGGCAATCCGAGGAGAGAGGTGCTGAGCTGTGATGCACTTGGGGAATTTC 1471
QY 1320 ----- 1319
DB 1472 CTGTGCGGAGAGCCAGGGCAAGAGAGTACGTGCTGTGGTGTGGATGCTG 1531
QY 1320 ----- 1319
DB 1532 CCCCCGACTTATCATCCAGGCTCACGCCCTCCAGAAATGAGGCTGTGACCCCCGGT 1591
QY 1320 -----GAACTGTATCCGATCTGGAAGGGGAGGCTTTTCTTACATTCCTTTCG 1366
DB 1592 CCCCCTGTGAGAACCTGTACCGACGTGAAGGGGAGGCTTTTCTTACATTCCTTTCG 1651
QY 1367 TCATGCAACCACTTACTGAGACCCAGAGCCCTCCACCAAGAGAGTGTGTGTCTG 1426
DB 1652 TCATGCAACCACTTACTGAGACCCAGAGCCCTCCACCAAGAGAGTGTGTGTCTG 1711
QY 1427 ACAGGCTGTGCAAGGACAAAGTGGTGTGAACCATGAGGACCTGTGTGGTGTGAC 1486
DB 1712 ACAGGCTGTGCAAGGACAAAGTGGTGTGAACCATGAGGACCTGTGTGGTGTGAC 1771
QY 1487 AGATTGAGAGGGGGAACCTTTGGCGAAGTGTCAAGCGGAAGCCCTGCGAGCGCAACACCC 1546
DB 1772 AGATTGAGAGGGGGAACCTTTGGCGAAGTGTCAAGCGGAAGCCCTGCGAGCGCAACACCC 1831
QY 1547 TGGTGGCGGTGAAGTCTTGTGAGAGACGCTCCACCTGACCTCAAGGCAAGTTCTTAC 1606
DB 1832 TGGTGGCGGTGAAGTCTTGTGAGAGACGCTCCACCTGACCTCAAGGCAAGTTCTTAC 1891
QY 1607 AGAAGCGAGATCTTGAAGAGTACAGCCACCCCAACATCGTGTCTATTTGTGTCT 1666
DB 1892 AGAAGCGAGATCTTGAAGAGTACAGCCACCCCAACATCGTGTCTATTTGTGTCT 1951
QY 1667 GCAACCAAGAGAGCCCATCTTACATGTGTATGAGAGTGTGCAAGGGGGCGACTTCTGA 1726
DB 1952 GCAACCAAGAGAGCCCATCTTACATGTGTATGAGAGTGTGCAAGGGGGCGACTTCTGA 2011
QY 1727 CTTTCTCCGCAAGGAGGGGCGCGCTGCGGGTGAAGCTCTGCTGAGATGATGATGGG 1786
DB 2012 CTTTCTCCGCAAGGAGGGGCGCGCTGCGGGTGAAGCTCTGCTGAGATGATGATGGG 2071
QY 1787 ATGCAAGCTGTGCAATGAGTACTTGAAGAGCAAGTGTGATCCAACGGGACCTTGGCTG 1846
DB 2072 ATGCAAGCTGTGCAATGAGTACTTGAAGAGCAAGTGTGATCCAACGGGACCTTGGCTG 2131
QY 1847 CTGCGAATGCTGTGTGACAGAGAAATGTCTGAAATCAATGATCTTTGGGATGTCCTC 1906
DB 2132 CTGCGAATGCTGTGTGACAGAGAAATGTCTGAAATCAATGATCTTTGGGATGTCCTC 2191
QY 1907 GAGAGGAGCGATGGGCTTATGCAAGGCTCAAGGGGCTCAAGCAAGTCCCGTGAAGT 1966
DB 2192 GAGAGGAGCGATGGGCTTATGCAAGGCTCAAGGGGCTCAAGCAAGTCCCGTGAAGT 2251
QY 1967 GAGCGCACTGAGGCTTTAACTACGGGCGCTTCTCTCGAAGCGAGGTGTGAGCT 2026

Db 2252 GAAACCCACTGAGCCCTTAATCTAC----- 2277
Qy 2027 TTGGCATCTTGCTCTGGAGACCTTCAGCCTGGGGCTCCCTCATCTCCAACTCAGCA 2086
Db 2278 ----- 2277
Qy 2087 ATCAGCAGACACGAGGAGTTTGTGAGAGAGGGGGCGCTCTGCTCTGCTCCAGAGCTGTTC 2146
Db 2278 -----GGGGGCGGTCTGCGCTCCAGAGCTGTTC 2308
Qy 2147 CTGATCCGCTGTTCAAGGCTCATGAGACAGTCTGAGCTATGAGCCTGGGACAGGCGCA 2206
Db 2309 CTGATCCGCTGTTCAAGGCTCATGAGACAGTCTGAGCTATGAGCCTGGGACAGGCGCA 2368
Qy 2207 GCTTCGACCATCTTACAGAGCTCAGAGCATCCGAAAGCGGCACTCG 2256
Db 2369 GCTTCGACCATCTTACAGAGCTCAGAGCATCCGAAAGCGGCACTCG 2418

RESULT 2
CR593957 1189 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0D1008YG07 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR593957
VERSION CR593957.1 GI:50474764
KEYWORDS HTC; cDNA
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1189)
AUTHORS Li, W.B., Gruber, C., Jeesee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalized
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1189)
REFERENCE Direct Submision
AUTHORS Genoscope.
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source 1. 1189
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1008YG07"
/issue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 44.4%; Score 1002; DB 3; Length 1189;
Best Local Similarity 100.0%; Pred. No. 1.2e-200;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1255 GGAAGAGACCCAGCTGAGATCTTTAAGAGCACATCTCAGAAATCTTCCGCCCAAG 1314
Db 1 GGAAGAGACCCAGCTGAGATCTTTAAGAGCACATCTCAGAAATCTTCCGCCCAAG 60
Qy 1315 TTCTCGAAGCTGTACGAGCTGAGAGGGGAAAGGCTTTCTAGACATCTCTTGTGCTCATCGAC 1374
Db 61 TTCTCGAAGCTGTACGAGCTGAGAGGGGAAAGGCTTTCTAGACATCTCTTGTGCTCATCGAC 120
Qy 1375 CACTTACTAGACACCAAGAGCCCTTCACCAAGAGAGTGTGTGTCTCGACAGGGCT 1434

Db 121 CACTTACTAGACACCAAGAGCCCTTCACCAAGAGAGTGTGTGTCTCTCGACAGGGCT 180
Qy 1435 GTGCCCAAGAGACAGTGGGTGCTGAACATGAGAGCATGTGTGTGGTGAAGCATTTGA 1494
Db 181 GTGCCCAAGAGACAGTGGGTGCTGAACATGAGAGCATGTGTGTGGTGAAGCATTTGA 240
Qy 1495 CGGGGGAATCTTGGGCAATGTTCAGCCGAGCCCTGCGAGCCGAGCAACACCTGTGTGGC 1554
Db 241 CGGGGGAATCTTGGGCAATGTTCAGCCGAGCCCTGCGAGCCGAGCAACACCTGTGTGGC 300
Qy 1555 GTGAAGCTTTGTGAGAGAGCGCTCCCACTGACCTCAAGGCAAGTTTCTACAGAAAGC 1614
Db 301 GTGAAGCTTTGTGAGAGAGCGCTCCCACTGACCTCAAGGCAAGTTTCTACAGAAAGC 360
Qy 1615 AGGATCTTGAAGACAGTACAGCAACCCCAATGTGTGTCTCATTTGTGTCTGACCCAG 1674
Db 361 AGGATCTTGAAGACAGTACAGCAACCCCAATGTGTGTCTCATTTGTGTCTGACCCAG 420
Qy 1675 AAGCAGCCCATCTTACATGCTCATGAGCTTTGTGCAAGGGGGGCACTTCTGACCTTCTC 1734
Db 421 AAGCAGCCCATCTTACATGCTCATGAGCTTTGTGCAAGGGGGGCACTTCTGACCTTCTC 480
Qy 1735 CGCAGGAGAGGGGGCCGCGCTGCGGAGAAAGTCTGCTGCAAGTGGTGGGAGTGCAGCT 1794
Db 481 CGCAGGAGAGGGGGCCGCGCTGCGGAGAAAGTCTGCTGCAAGTGGTGGGAGTGCAGCT 540
Qy 1795 GCTGGCAGTGAAGTACCTGAGAGCAAGTCTGATCACCAGGACCTGTGCTGGAAAC 1854
Db 541 GCTGGCAGTGAAGTACCTGAGAGCAAGTCTGATCACCAGGACCTGTGCTGGAAAC 600
Qy 1855 TGCTGTGACAGAGAAAGTGTCTGAGATCAGTGAATCTTGGATGTCCGAGAGAA 1914
Db 601 TGCTGTGACAGAGAAAGTGTCTGAGATCAGTGAATCTTGGATGTCCGAGAGAA 660
Qy 1915 GCCGATGGGGTCTATGACAGCTCAGGGGGCTCAGACAGTCCCGTGAAGTGCAGCGCA 1974
Db 661 GCCGATGGGGTCTATGACAGCTCAGGGGGCTCAGACAGTCCCGTGAAGTGCAGCGCA 720
Qy 1975 CCTGAGGCCCTTAATACAGGCGGCTAATCTCTCGAAAGCGAGCTGTGAGCTTGGCATC 2034
Db 721 CCTGAGGCCCTTAATACAGGCGGCTAATCTCTCGAAAGCGAGCTGTGAGCTTGGCATC 780
Qy 2035 TTGCTTGGGAGACCTTACGCTGCGGGGCTCTCCCTATTCCTCAAGCTTCAAGCATCAGC 2094
Db 781 TTGCTTGGGAGACCTTACGCTGCGGGGCTCTCCCTATTCCTCAAGCTTCAAGCATCAGC 840
Qy 2095 AACCGGAGTTTGTGAGAGAGGGGGCGCTGCGCTGCGGCAAGGCTGTCTGATGCC 2154
Db 841 AACCGGAGTTTGTGAGAGAGGGGGCGCTGCGCTGCGGCAAGGCTGTCTGATGCC 900
Qy 2155 GTGTTGAGCTCATGAGAGAGTGTGCGGCTTATGAGCTTGGGAGCGGCCCATCTTCAAGC 2214
Db 901 GTGTTGAGCTCATGAGAGAGTGTGCGGCTTATGAGCTTGGGAGCGGCCCATCTTCAAGC 960
Qy 2215 ACCATTTACAGAGCTTCAAGACATCCGAAAGCGGCAATCGG 2256
Db 961 ACCATTTACAGAGCTTCAAGACATCCGAAAGCGGCAATCGG 1002

RESULT 3
BX356089 936 bp mRNA linear EST 08-APR-2004
LOCUS BX356089 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION BX356089 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
Accession BX356089
Version BX356089.2 GI:46290249
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 936)

AUTHORS Li, W.B., Gruber, C., Jeesee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On May 5, 2003 this sequence version replaced gi:30372027.

Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 7663.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CSOD1008AD04P1ac=7663.f.

FEATURES

source

Location/Qualifiers

1..936
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSOD1008Y07"
 /library="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 40.9%; Score 923.6; DB 5; Length 936;
 Best Local Similarity 99.8%; Pred. No. 4.1e-184;
 Matches 934; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1255 GGAGAGACACCCACGCTGAGATCTTAAAGACCAATCTTCGCCCCCAAG 1314
 DB 1 GGAGAGACACCCACGCTGAGATCTTAAAGACCAATCTTCGCCCCCAAG 60
 QY 1315 TTCTCGAACCTTGACGAGTGAAGAGGGAAGGCTTCTCGAATCTTCGATCGAC 1374
 DB 61 TTCTCGAACCTTGACGAGTGAAGAGGGAAGGCTTCTCGAATCTTCGATCGAC 120
 QY 1375 CACCTAAGAGACCCAGACGACCCCTCACCAGAAAGTGTGTTGCTCGACAGGCT 1434
 DB 121 CACCTAAGAGACCCAGACGACCCCTCACCAGAAAGTGTGTTGCTCGACAGGCT 180
 QY 1435 GTGCCCAAGACAAAGTGGTGTGAACCATGAGAGACTGTGTGGGTAGAGATTGGA 1494
 DB 181 GTGCCCAAGACAAAGTGGTGTGAACCATGAGAGACTGTGTGGGTAGAGATTGGA 240
 QY 1495 CGGGGGAACCTTGGCGAAGTTCAGCGAGCGCTGCGAGCCGACCAACCTCGTGGCG 1554
 DB 241 CGGGGGAACCTTGGCGAAGTTCAGCGAGCGCTGCGAGCCGACCAACCTCGTGGCG 300
 QY 1555 GTGAAGCTTGTGAGAGAGCTCCACCTGACCTCAAGGCCAAGTTTCTACAGAGCG 1614
 DB 301 GTGAAGCTTGTGAGAGAGCTCCACCTGACCTCAAGGCCAAGTTTCTACAGAGCG 360
 QY 1615 AGAATCTTGAAGACGTAACGACCCCAACATGTGCTTCAATGGTGTCTGACCCAG 1674
 DB 361 AGAATCTTGAAGACGTAACGACCCCAACATGTGCTTCAATGGTGTCTGACCCAG 420
 QY 1675 AAGCAGCCATCTACATGTCGAGAGCTGTGCAAGGGGGGCACTTCCGACCTTCCTC 1734
 DB 421 AAGCAGCCATCTACATGTCGAGAGCTGTGCAAGGGGGGCACTTCCGACCTTCCTC 480
 QY 1735 CGCAGCGAGGGGCGCGCTGCGGGTGAAGACTCTGCTGAGATGTGGGGATGAGCT 1794
 DB 481 CGCAGCGAGGGGCGCGCTGCGGGTGAAGACTCTGCTGAGATGTGGGGATGAGCT 540
 QY 1795 GCTGGCATGAGATCTGTAAGAGCAAGTGTGATCCACCGGACCTGCTGCTCGAAC 1854
 DB 541 GCTGGCATGAGATCTGTAAGAGCAAGTGTGATCCACCGGACCTGCTGCTCGAAC 600

QY 1855 TGCCTGTGACAGAGAGATGTCTGAAGATCAGTACTTTGGATGTCCGAGAGAA 1914
 DB 601 TGCCTGTGACAGAGAGATGTCTGAAGATCAGTACTTTGGATGTCCGAGAGAA 660
 QY 1915 GCCGATGGGGTTTATGACCTTCAGGGGGCTTCAGACAGTCCCTGAAGTGAACCGA 1974
 DB 661 GCCGATGGGGTTTATGACCTTCAGGGGGCTTCAGACAGTCCCTGAAGTGAACCGA 720
 QY 1975 CCTGAGGCCCTTAATCAGCGCGCTACTCTCCGAAGGAGGTGAGACTTTGGCATC 2034
 DB 721 CCTGAGGCCCTTAATCAGCGCGCTACTCTCCGAAGGAGGTGAGACTTTGGCATC 780
 QY 2035 TTGCTCTGGAGACCTTCAGCTT-GGGGGCTTCCCTTAATCCCACTCAGCAATCAGA 2093
 DB 781 TTGCTCTGGAGACCTTCAGCTT-GGGGGCTTCCCTTAATCCCACTCAGCAATCAGA 840
 QY 2094 GACACGGAGATTGTGAGAGAGGGGGCGCTGTCCCTCCAGAGCTGTGCTGATGC 2153
 DB 841 GACACGGAGATTGTGAGAGAGGGGGCGCTGTCCCTCCAGAGCTGTGCTGATGC 900
 QY 2154 CGTGTTCAGGCTCATGAGACAGTCTGGGCTTATGA 2189
 DB 901 CGTGTTCAGGCTCATGAGACAGTCTGGGCTTATGA 936
 RESULT 4
 AK004587
 LOCUS 3373 bp mRNA linear HTC 03-APR-2004
 DEFINITION Mus musculus adult male lung cDNA, RIKEN full-length enriched
 library, clone:1200003015 product:similar to PROTO-ONCOGENE
 TYROSINE-PROTEIN KINASE FES/FPS (BC 2.7.1.112) (C-FES) [Mus
 musculus], full insert sequence.
 AK004587
 ACCESSION
 VERSION AK004587.1 GI:12835858
 KEYWORDS HTC; CAP trapper
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multiplexed capillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE 5

AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 3373)

AUTHORS Adachi, T., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carinetti, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komoh, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saio, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, S., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toy, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT Please visit our web site (http://genome.gsc.riken.jp/) for further details.

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer 15' GAGGAGAGACGGCCGCGACTCGACTGATTTTATTTTATTTT 3', cDNA was prepared by using retroviral thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGGAGAGAGATCCCAAGCTCAATTTATTTTAAACCCCCCCC 3'. cDNA was cleaved with XhoI and SbfI. Cloning sites, 5' end: SbfI; 3' end: XhoI. Host: SOLR.

FEATURES Location/Qualifiers

source 1. 3373

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM:DB:1200003015"

/db_xref="taxon:10090"

/clone="1200003015"

/sex="male"

/issue_type="lung"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

1. 3373

/note="similar to PROTO-ONCOGENE TYROSINE-PROTEIN KINASE PPS/PPS (EC 2.7.1.112) (C-FPS) [mus musculus] (SWISSPROT:P16879, evidence: PASTY, 77.2%id, 100%length, match=3083)"

misc_feature

ORIGIN

Query Match 39.2%; Score 885.4; DB 3; Length 3373;

Best Local Similarity 77.8%; Pred. No. 5.6e-176;

Matches 1146; Conservative 0; Mismatches 191; Indels 136; Gaps 1;

1 ATGGGCTTCTTCTGAGCTGTGACAGCCCGACAGGCGAGGCTCTGACGAATATGAC 60

112 ATGGGCTTCTTCTGAGCTGTGACAGCCCGACAGGCGAGGCTCTGACGAATATGAC 171

61 GAGGCGGAGCTTCTGCTGAGGAGGAGGAGAAAGTGAATGAGGCGGAGTCAAGAGT 120

172 GAGGCTGAGCTGCTGCTGAGGAGGAGGAGGAGAAATGAGTGGCCCAAGAGGCTCAAGAGT 231

121 GACAGGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180

232 GACCGGGAAATATGACAGATTTCTTACCAATATCTCTTGACAGGACCGAGGCCAGAGC 291

181 CCGGCGCATAGCCCTTGACAGCCCGCATAGTCAATCTGAGGCTGAGATCACAGCCAAACT 240

292 TGGAGCAGCGGCCCGACAGCCCTGACAGGCTCTGGGGCAGAGATTAACAGCCAGAGACC 351

241 GAGGCGCTAGCGGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300

352 GAGAACTTAGCGCGGCTGCTGGGCGAGATGAGAAATCTGAATCTGGGGCCCTTGAGC 411

301 AAGCTGAGCTGCTCTCCGGGAAACCGGAGAGCTTCCGAAAGCTTACAGAGCAGAGCTGG 360

412 AAATGAGCTGCTGCTCTCCGGGAGCGGAGAGCTTGAAGAAAGCTTACAGAGCAGAGCTGG 471

361 CAGCAGCTGACAGGAGGAGTCAACCAAGCCACAGCCAGGAGCATTGAGAGCTGAAGAGC 420

472 CAGCAGCTGACAGGAGGAGTCAACCAAGCCACAGCCAGGAGCATTGAGAGCTGAAGAGC 531

421 CAGTACCGAGCTCTGACAGGAGCAGTGGCCAAAGCCAAAGCTTACAGAGGAGGAGC 480

532 CAGTACCGAGCTCTGATGAGATGACAGCCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGC 591

481 AAAGACAGAGCCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540

592 AAAGACAGAGCCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 651

541 GCTCAGCAGACCGCTATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600

652 GCCACACACACCGCTATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 711

601 CACGAGCTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660

712 CACGAGCTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 771

661 TGCATCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720

772 GGCATCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 831

721 GTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780

832 GTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 891

781 TACCAAGGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840

892 TACCTAGGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 951

841 GATGAGTCACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900

952 GATGAGTCACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1011

901 CTGACTGTGAGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 926

1012 TTGACCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1071

927 -----

1072 GTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1131

927 -----

927 -----

1132 TGGTGGGGGCGCAGCTTTTATTTCTCTGCGCTTGGGGGATCAGAGCTGATCTTGTGACAGA 1191

945 TGAAGCTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1004

1192 TGAAGCTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1251

1005 ACAAGAGCTCGGAAATGAAGAAGAAACCCACCCCGGAGCGGGGTGAGCTGTGGG 1064

1252 GCGTGAAGCTCAAAAGTGAAGAAACAGAAACCCACCCCGGAGCGGGGTGAGCTGTGGG 1311

1065 CAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1124

1312 CAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1371

QY 1125 CAAGTCGAGGCGCCAGAGAGTTCCTGAGACCAAGCTGAGACCTTGAGGCGCCCGGCA 1184
DB 1372 CAGGCTTCAGGCGCCAGAGACTTCTGAGAGCAAGATGAGACCTGCGGCGGCA 1431
QY 1185 GCGCCCGCTGTCCTGTCCTGAGAGATGACCGGCACTCCAGTCCTCTCGAGAGAGA 1244
DB 1432 GCGCCAGCGTCGCTCTTCAAGATGACCGGCACTCCAGTCCTCTCGAGAGAGA 1491
QY 1245 GCGAGAGGAGAGAGAGACCCGCTGAGAGATCTTAAGAGCAATCTCAGAGATCTT 1304
DB 1492 GCGAGAGGAGAGAGAGACCCGCTGAGAGATCTTAAGAGCAATCTTCCGAGATCTT 1551
QY 1305 CCGCCCGAGTTCGACACCTGACAGGACTGGA 1337
DB 1552 CCGCCCGAGTTCGACACCTGACAGGACTGGA 1584

RESULT 5
AL542976 936 bp mRNA linear EST 24-MAR-2004
LOCUS AL542976 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE0131P02
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION AL542976 GI:45718543
VERSION AL542976.3 GI:45718543
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 936)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:30548660.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 7663.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0DE013D101P1&c=7663.f.
FEATURES
source
1..936
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE0131P02"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 37.6%; Score 848.8; DB 1; Length 936;
Best Local Similarity 99.2%; Pred. No. 2.5e-168;
Matches 861; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 ATGGGCTTCTCTTCTGAGCTGTCAGGCGCCAGGCGGCTCTGACGAAATGACG 60
DB 57 ATGGGCTTCTCTTCTGAGCTGTCAGGCGCCAGGCGGCTCTGACGAAATGACG 116
QY 61 GAGGCGGAGCTTCTCTTCTGAGGCGCATAGAAAGTGATGCGCCAGCGGCTCAAGAGT 120
DB 117 GAGGCGGAGCTTCTCTTCTGAGGCGCATAGAAAGTGATGCGCCAGCGGCTCAAGAGT 176

QY 121 GACAGGAGATATGACAGGACTGCTTCAACCATATGCTCTTCAAGAGAGTGGGGCCAGAGC 180
DB 177 GACAGGAGATATGACAGGACTGCTTCAACCATATGCTCTTCAAGAGAGTGGGGCCAGAGC 236
QY 181 GGGGCGCATATGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 237 GGGGCGCATATGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 296
QY 241 GAGGCGCATATGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 297 GAGGCGCATATGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 356
QY 301 AAGCTGAGCTGCTGATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 357 AAGCTGAGCTGCTGATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 416
QY 361 CAGCAGCTGACAG 420
DB 417 CAGCAGCTGACAG 476
QY 421 CAGTACCGAGCTTGGACAG 480
DB 477 CAGTACCGAGCTTGGACAG 536
QY 481 AAAG 540
DB 537 AAAG 595
QY 541 GCTTACACAG 600
DB 596 GCTTACACAG 655
QY 601 CAGCAGCTTCTGCTGCGGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGG 660
DB 656 CAGCAGCTTCTGCTGCGGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGG 715
QY 661 TGCATCTTGAAG 720
DB 716 TGCATCTTGAAG 775
QY 721 GGTGTCGATTCACCGGAG 780
DB 776 GGTGTCGATTCACCGGAG 835
QY 781 TGCAGAGCTTCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGG 840
DB 836 TGCAGAGCTTCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGG 895
QY 841 GATGAGCTCATGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 868
DB 896 GATGAGCTCATGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 923

RESULT 6
BM456755 1029 bp mRNA linear EST 05-FEB-2002
LOCUS BM456755
DEFINITION AGENDOURT 6403957 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5583652
5', mRNA sequence.
ACCESSION BM456755 GI:18505795
VERSION BM456755
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1029)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Prepared by: Life Technologies, Inc.
DNA Sequencing by: I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>

FEATURES

Bouyce

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5583652"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="N1H.MGC.92"
/note="Organ: testis; Vector: pCMV.SPORT6; Site 1: NotI, Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies Note: this is a N1H.MGC library."

```

ORIGIN

Query Match	35.7%;	Score 805.8;	DB 4;	Length 1029;
Best Local Similarity	97.9%;	Pred. No. 3e-159;		
Matches 869; Conservative	0;	Mismatches 13;	Indels 6;	Gaps 5

QY	1	ATGGGCTCTCTTCTGAGCTGTGCAGCCCCCAGGGGCAAGGGGTCCTGTGACAAATGACG	60
Db	81	ATGGGCTCTCTTCTGAGCTGTGCAGCCCCCAGGGGCAAGGGGTCCTGTGACAAATGACG	140
QY	61	GAGGCCGAGCTTCGTCTACTGGAGGACATGAGAAAGTGGATGACCGGTCAAAGAT	120
Db	141	GAGGCCGAGCTTCGTCTACTGGAGGACATGAGAAAGTGGATGACCGGTCAAAGAT	200
QY	121	GACAGGGAGTATGACAGGACCTGGTTACCAACATGTCCCTGCACAGACAGTGGGGGCCAGAGC	180
Db	201	GACAGGGAGTATGACAGGACCTGGTTACCAACATGTCCCTGCACAGACAGTGGGGGCCAGAGC	260
QY	181	CGGGCCATCAGCCCTTGACAGCCCCATCATGTCAGTCCTTGGGCTGAGATTCACCAAGCCAAACT	240
Db	261	CGGGCCATCAGCCCTTGACAGCCCCATCATGTCAGTCCTTGGGCTGAGATTCACCAAGCCAAACT	320
QY	241	GAGGGCTCTGAGCCCGCTTGCTGTGGGACACACGACAGAGATCTGAACTCAGGGCCCTCTGAC	300
Db	321	GAGGGCTCTGAGCCCGCTTGCTGTGGGACACACGACAGAGATCTGAACTCAGGGCCCTCTGAC	380
QY	301	AAGCTGAGCTGTGTCATCCGGGAAACGGACAGAGCTTCGAAAGACTCAACAGCAGCAGTGG	360
Db	381	AAGCTGAGCTGTGTCATCCGGGAAACGGACAGAGCTTCGAAAGACTCAACAGCAGCAGTGG	440
QY	361	CAGCAGCTGCACAGAGAGCTTCCAAAGACCCACAGCCAGACATTTGAGAGCTGAAGAC	420
Db	441	CAGCAGCTGCACAGAGAGCTTCCAAAGACCCACAGCCAGACATTTGAGAGCTGAAGAC	500
QY	421	CAGTACCGAGCTGTGGCAACGGGACAGTGCCCAAGCCAAAGCCAAAGTACCAAGAGGCGACG	480
Db	501	CAGTACCGAGCTGTGGCAACGGGACAGTGCCCAAGCCAAAGCCAAAGTACCAAGAGGCGACG	560
QY	481	AAAGACAAAGAACCGTGCACAGGACCAAGGACAAAGTATGTGCGCAGCCTGTGGAAGCTCTTT	540
Db	561	AAAGACAAAGAACCGTGCACAGGACCAAGGACAAAGTATGTGCGCAGCCTGTGGAAGCTCTTT	620
QY	541	GCTCACCAACACCGCTATGTGCTGGGCGTGGCGGGCTGGCCAGCTACACCAACAGCACAC	600
Db	621	GCTCACCAACACCGCTATGTGCTGGGCGTGGCGGGCTGGCCAGCTACACCAACAGCACAC	680
QY	601	CACCAAGCTCTGCTCACCAGGCTGCTGCGGTCAGTGCACAGACCTTGACAGAGAGATGGCT	660
Db	681	CACCAAGCTCTGCTCACCAGGCTGCTGCGGTCAGTGCACAGACCTTGACAGAGAGATGGCT	740
QY	661	TGCATCTGAAAGAGATCCTGCAGAGATACCTTGAGATTAGCAGCCTGTGTGCAGATGAG	720

Accession	Sequence	Position
Db	TGCATCCGAAAGAGATCTGSCAGSAATACCTGGAGATTAGCAACTGTGGTCAGGATGAG	800
Qy	GTGCTGSCCATTTCAACCGG-AGATGCTGTGCAAGCTGTGCTCCCGCATTCCAAGCTGAAGCTTA	779
Db	GTGTGTGGCATTTCAACCGGAGATGTGGCTGCAAGCTGTGTCTCCGCATTCCAAGCTGAGGCTTA	860
Qy	GTACCAAGGCTCTCTGCGACAGTATGGTCTCCGCACTTAC-GTCCACCTCTGTGTCAG-837	
Db	GTACCAAGGCTT-CTGGCAGAGTATGGTCTCCGCAACTGACAGGTTCACACCTGTGTGTCAGT-919	
Qy	TTGCATGATGACTGCTTGA--GAGAGGTGAACCGCTGAGGCTTGGGG-883	
Db	TTGCATGATGACTGCTTGAAGGAGAGGGGTGAACCGCTGGAACCTTGGGG-967	

RESULT 7				
LOCUS	BX359010	968 bp	mRNA	linear
DEFINITION	BX359010 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CSODID051Y020 5-PRIME, mRNA sequence.			
ACCESSION	BX359010			
VERSION	BX359010.2	GI:46307488		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			

REFERENCE 1 (bases 1 to 368)
AUTHORS Li, W. B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 5, 2003 this sequence version replaced gi:30378289.

For more information about this cluster, see <http://www.genoscope.cns.fr/cdna?c=CS0D1051B110QPlc=7663.f>

FEATURES

Source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/md_xref="taxon:9606"
/clone="CS01051Y020"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and cloned into the NotI and EcoR V
sites of the pCMSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	34.7%	Score 783.6;	DB 5;	Length 968;
Best Local Similarity	95.9%	Pred. No. 1.4e-154;		
Matches 879;	Conservative	9;	Mismatches 20;	Indels 9;
				Gaps 8

QY 1 ATGGGCTTCCTTCGACCTGTGACGCCCCCAGGGCCACGGGGTCTGTGACGAAATGCG 60
Db 61 ATGGGCTTCCTTCGACCTGTGACGCCCCCAGGGCCACGGGGTCTGTGACGAAATGCG 120
QY 61 GAGGCCGAGCTTCGTCTACTGAGGCGCATGAGAAATGAGATGGCCACGGCGGTCAAGAT 120
Db 121 GAGGCCGAGCTTCGTCTACTGAGGCGCATGAGAAATGAGATGGCCACGGCGGTCAAGAT 180
QY 121 GACACGGAGTATGACAGACTGCTTACCACATGTCCCTGCAAGACAGTGGGGGCCAGAAC 180

```

Db      181 GACGAGGAGTATGACAGAGCTGCTTCAACAATGTCCTCTGACGAGCAGTGGGGCCAGAGC 240
QY      181 |CGGGCCATCAGCCCTGACAGCCCATCAGTCACTGCTGGCTGAGATCAACAGCCAACT 240
Db      241 CGGGCCATCAGCCCTGACAGCCCATCAGTCACTGCTGGCTGAGATCAACAGCCAACT 300
QY      241 GAGGGCTGAGCCGCTTGTGCTGGGCGACAGCAGAGGATCTGAATCTCAGGGCCCTGAGC 300
Db      301 GAGGGCTGAGCCGCTTGTGCTGGGCGACAGCAGAGGATCTGAATCTCAGGGCCCTGAGC 360
QY      301 AAGCTGAGCTGCTCATCCGGGAAACGGCAGACGCTTGGCAAGCTTACAGCGACGCTGG 360
Db      361 AAGCTGAGCTGCTCATCCGGGAAACGGCAGACGCTTGGCAAGCTTACAGCGACGCTGG 420
QY      361 CAGCAGCTGACAGAGAGCTCACCAAGACCACAGCAGAGCATTGGAAGAGTGAAGAGC 420
Db      421 CAGCAGCTGACAGAGAGCTCACCAAGACCACAGCAGAGCATTGGAAGAGTGAAGAGC 480
QY      421 CAGTACCGAGCTTGGCAGCGGACAGTGCACCAAGCCAAAGCCAAAGTACAGAGGCGCAGC 480
Db      481 CAGTCCGAGCTTGGCAGCGGACAGTGCACCAAGCCAAAGCCAAAGTACAGAGGCGCAGC 540
QY      481 AAGAGCAAGAGACCTGTGACCAAGGCCAAAGACAAATATGTGGCAGCCTGTGGAAGCTCTT 540
Db      541 AAGAGCAAGAGACCTGTGACCAAGGCCAAAGACAAATATGTGGCAGCCTGTGGAAGCTCTT 599
QY      541 GCTCACCAACAGCGCTATGTGTGGGGGCTGGGGCTGGCGAGCTACACCAAGCAGCAGC 600
Db      600 GCTCACCAACAGCGCTATGTGTGGGGGCTGGGGCTGGCGAGCTACACCAAGCAGCAGC 657
QY      601 CACCAAGCTCTGTGCTGGCCGCTGCTGCGGTCACTGACAGACCTGACAGAGAGATGGCT 660
Db      658 CACCAAGCTCTGTGCTGGCCGCTGCTGCGGTCACTGACAGACCTGACAGAGAGATGGCT 717
QY      661 TGCATCTTGAAGAGATCTCTGACGAATATCTGGAATTTAGCACTGTGTGACGATGAG 720
Db      718 TGCATCTTGAAGAGATCTCTGACGAATATCTGGAATTTAGCACTGTGTGACGATGAG 777
QY      721 GTGGTGGCCATTACCGGGAGATGAGTGGCTGAGCTGCTGCCCGCATCCAGCCTGAGGCTGAG 780
Db      778 GTGGTGGCCATTACCGGGAGATGAGTGGCTGAGCTGCTGCCCGCATCCAGCCTGAGGCTGAG 834
QY      781 TACCAAGGCTTCTGCGACAGTATGAGTCCGACCTGACGCTCCACCTGTGTCAAGTTT 840
Db      835 TA-CAAGGCTTCTGCGACAGTATGAGTCCGACCTGACGCTCCACCTGTGTCAAGTTT 892
QY      841 GATGAGTCACTGCTTGAAGAGGAGTGAACCGCTGAGAGCTTCCAGCTGAACGAG 900
Db      893 GATGAGTCACTGCTTGAAGAGGAGTGAACCGCTGAGAGCTTCCAGCTGAACGAG 951
QY      901 CTGACTGTGAGAGCGT 917
Db      952 CTGACTGTGAGAGAGT 968

```

```

RESULT 8
BX381726 939 bp mRNA linear EST 26-Apr-2004
LOCUS BX381726 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS001069YK21 5-PRIME, mRNA sequence.
ACCESSION BX381726
VERSION BX381726.2 GI:46570766
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE 1
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLES (Base 1 to 939)
JOURNAL Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
COMMENT Unpublished (2001)
Full-length cDNA libraries and normalization
On May 8, 2003 this sequence version replaced gi:30443719.
Contact: Genoscope

```

```

FEATURES
source
1..939
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS001069YK21"
/cisue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the Not I and EcoR V sites of the pcMVSPORT 6 vector. Library was normalized."
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?as=CS001069AF10P1ac=7663.f.
Location/Qualifiers

```

ORIGIN

```

Query Match 34.3%; Score 773.4; DB 5; Length 939;
Best Local Similarity 97.2%; Pred. No. 2e-152;
Matches 845; Conservative 13; Mismatches 4; Indels 7; Gaps 7;

```

```

QY      1 ATGGGCTTCTCTTCTGAGCTGTGACAGCCCGCAGGCGCCAGCGGGTCTGTGACCAATGACG 60
Db      32 ATGGGCTTCTCTTCTGAGCTGTGACAGCCCGCAGGCGCCAGCGGGTCTGTGACCAATGACG 91
QY      61 GAGGCGGAGCTTCTGTATGAGAGGCGATGAGAAATGATGGCCAGCGGGTCAAGAT 120
Db      92 GAGGCGGAGCTTCTGTATGAGAGGCGATGAGAAATGATGGCCAGCGGGTCAAGAT 151
QY      121 GACAGGAGTATGACAGAGCTGCTTACCAATGTCCCTGACAGACAGTGGGGCCAGAGC 180
Db      152 GACAGGAGTATGACAGAGCTGCTTACCAATGTCCCTGACAGACAGTGGGGCCAGAGC 211
QY      181 CGGGCCATCAGCCCTGACAGCCCATCAGTCACTGCTGGCTGAGATTCACAGCCAACT 240
Db      212 CGGGCCATCAGCCCTGACAGCCCATCAGTCACTGCTGGCTGAGATTCACAGCCAACT 271
QY      241 GAGGGCTGAGCGCTTGTGCGGAGCAAGGAGAGATCTGAATCTCAGGGCCCTGAGC 300
Db      272 GAGGGCTGAGCGCTTGTGCGGAGCAAGGAGAGATCTGAATCTCAGGGCCCTGAGC 331
QY      301 AAGCTGAGCTGCTCATCCGGGAAACGGCAGACGCTTGGCAAGCTTACAGCGAGCAGTGG 360
Db      332 AAGCTGAGCTGCTCATCCGGGAAACGGCAGACGCTTGGCAAGCTTACAGCGAGCAGTGG 391
QY      361 CAGCAGCTGACAGAGAGCTCACCAAGACCACAGCAGAGCATTGGAAGAGTGAAGAGC 420
Db      392 CAGCAGCTGACAGAGAGCTCACCAAGACCACAGCAGAGCATTGGAAGAGTGAAGAGC 451
QY      421 CAGTACCGAGCTTGGCAGCGGACAGTGCACCAAGCCAAAGGCGCAAGTACAGAGAGGCGCAGC 480
Db      452 CAGTACCGAGCTTGGCAGCGGACAGTGCACCAAGCCAAAGGCGCAAGTACAGAGAGGCGCAGC 511
QY      481 AAGAGCAAGAGACCTGTGACCAAGGCCAAAGACAAATATGTGGCAGCCTGTGGAAGCTCTT 540
Db      512 AAGAGCAAGAGACCTGTGACCAAGGCCAAAGACAAATATGTGGCAGCCTGTGGAAGCTCTT 570
QY      541 GCTCACCAACAGCGCTATGTGTGGGGCTGGGGCTGGCGAGCTTACCAACAGCAGCAGC 600
Db      571 GCTCACCAACAGCGCTATGTGTGGGGCTGGGGCTGGCGAGCTTACCAACAGCAGCAGC 629
QY      601 -CACCAAGCTCTGCTGCGCGGCTGTGTGGGTCACTGACAGAGCTTCAAGAGAGATGGC 659
Db      630 ACACAGCTCTGCTGCGCGGCTGTGTGGGTCACTGACAGAGCTTCAAGAGAGATGGC 689

```

QY 660 TTGCATCTGAAGAGATCTGCGAGGAATACTGAGATTAGAGAGCTGGTGCAGATGA 719
DB 690 TTGCATCTGGAAGAGATCTGCGAGGAATACTGAGATTAGAGAGCTGGTGCAGATGA 749
QY 720 GGTGGTGGCATTACCGGAGATGCTGCGAGCTGCTGCCGATCCAGCTGAGGCTGA 779
DB 750 GGTGGTGGCATTACCGGAGATGCTGCGAGCTGCTGCCGATCCAGCTGAGGCTGA 809
QY 780 GTCACAGGCTTCTGCGAGATGATGGTCCGACCTGACCTCCACCTGTGTCACTT 839
DB 810 GTCACAGGCTTCTGCGAGATGATGGTCCGACCTGACCTCCACCTGTGTCACTT 866
QY 840 CGATGATCACTGCTGAGGAGGCTGAC 868
DB 867 CGATG-RKTAATGCTTGAAGAGGCTGAC 894

RESULT 9
BO708270 948 bp mRNA linear EST 16-JUL-2002
LOCUS AGENCOURT_8484911 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6301278
DEFINITION 5', mRNA sequence.
ACCESSION BO708270
VERSION BO708270.1 GI:21847169
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 948)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Dr. Mark Watson
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM2516 row: 0 column: 07
High quality sequence stop: 610.
Location/Qualifiers
1. 948
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6301278"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_113"
/note="Organ: spleen; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN
Query Match 34.2%; Score 771.6; DB 5; Length 948;
Best Local Similarity 95.3%; Pred. No. 4.8e-152;
Matches 818; Conservative 0; Mismatches 34; Indels 6; Gaps 2;
QY 223 GAGATACCAAGCAAGTGAAGGCGCTGAGCCGCTTCTGCGGAGAGACGAGAGATCTG 282
DB 1 GAGATACCAAGCAAGTGAAGGCGCTGAGCCGCTTCTGCGGAGAGACGAGAGATCTG 60
QY 283 AACTCAGGCGCCCTGAGCAAGTGAAGCTGCTCATCCGGAACGAGCAGAGCTTCCCAAG 342
DB 61 AACTCAGGCGCCCTGAGCAAGTGAAGCTGCTCATCCGGAACGAGCAGAGCTTCCCAAG 120

QY 343 ACCTAAGGAGAGAGTGGAGAGAGCTGACAGAGAGACTCAACCAAGACCCAGACAGAC 402
DB 121 ACCTAAGGAGAGAGTGGAGAGAGCTGACAGAGAGACTCAACCAAGACCCAGACAGAC 180
QY 403 ATTGAAGAGCTGAAGAGCAGTACCGAGCTTGGACAGGGAACAGTCCCAAGCAAGCGC 462
DB 181 ATTGAAGAGCTGAAGAGCAGTACCGAGCTTGGACAGGGAACAGTCCCAAGCAAGCGC 240
QY 463 AAGTACCAAGAGGCGCAAGCAAGCAAGACCGTGAACAGGCGCAAGCAAGATATGTCGC 522
DB 241 AAGTACCAAGAGGCGCAAGCAAGCAAGACCGTGAACAGGCGCAAGCAAGATATGTCGC 300
QY 523 AGCTGTGGAAGCTCTTGTGCTACCAACCGGTATGTGTGTGGCGCTGGCGCTGGCGAG 582
DB 301 AGCTGTGGAAGCTCTTGTGCTACCAACCGGTATGTGTGTGGCGCTGGCGCTGGCGAG 360
QY 583 CTACACCAAGAGAGCAGACCAAGCAGCTGCTGCTGCTGGCGCTGGCGCTGCTGAGAGAC 642
DB 361 CTACACCAAGAGAGCAGACCAAGCAGCTGCTGCTGCTGGCGCTGGCGCTGCTGAGAGAC 420
QY 643 CTGACAGAGAGATGAGCTTGCATCTGAAAGAGATCTGCAAGATACCTGAGATTAGC 702
DB 421 CTGACAGAGAGATGAGCTTGCATCTGAAAGAGATCTGCAAGATACCTGAGATTAGC 480
QY 703 AGCTGTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 762
DB 481 AGCTGTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 763 ATCCAGCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 822
DB 541 ATCCAGCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 823 CCAACCTGTGTGTCAGTTCGATGATGATGATGATGATGATGATGATGATGATGATG 882
DB 601 CCAACCTGTGTGTCAGTTCGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 883 GAGCTTCAGCTGACAG 942
DB 661 GAGCTTCAGCTGACAG 720
QY 943 GATGAGCTGAGCTGTGCGACCGAGATGATGATGATGATGATGATGATGATGATGATG 1000
DB 721 GATGAGCTGAGCTGTGCGACCGAGATGATGATGATGATGATGATGATGATGATGATG 780
QY 1001 TGCAACAGAGAGCTCCGAAATGAAGAGA----GAACACCAACCCCGGAGAGCGGTG 1056
DB 781 TGCAACAGAGAGAGCTCCGAAATGAAGAGA----GAACACCAACCCCGGAGAGCGGTG 840
QY 1057 CTGCTGGGCAAGAGGCA 1074
DB 841 CAGCCTGCTGGGCGCA 858

RESULT 10
BX356088/c 980 bp mRNA linear EST 08-APR-2004
LOCUS BX356088 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone GS0D1008Y07 3-PRIME, mRNA sequence.
ACCESSION BX356088
KEYWORDS BX356088.2 GI:46288403
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 980)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:3036992.
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage

OY	560	TGCTGGGGGTGAGGGCTGCGCAGCTACACAC	CACACACACACACACAGCTCCCTGCGCCG	619
Db	249	TGCTGGGGGTGCGGGCTCGCAGCTACACAC	CACACACACACACACAGCTCTCTGCTGGCCG	308
OY	620	GCTTGCTCGGTCACCTGCAAGACCTTGCA	CGAGAGATGGCTTGCATCTCTGAAGAGATCC	679
Db	309	GCTTGCTCGGTCACCTGCAAGACCTTGCA	CGAGAGATGGCTTGCATCTCTGAAGAGATCC	368
OY	680	TGCAGGAATACCTGGAGATTGAGACCTGGAG	CACAGAGTGGTGGCCATTCCACCGGG	739
Db	369	TGCAGGAATACCTGGAGATTGAGACCTGGAG	CACAGAGTGGTGGCCATTCCACCGGG	428
OY	740	AGATGGCTGACAGCTGCTGCCCGCATCCA	GCTGAGGCTGAGTACCAAGGCTTCCCTGCGAC	799
Db	429	AGATGGCTGACAGCTGCTGCCCGCATCCA	GCTGAGGCTGAGTACCAAGGCTTCCCTGCGAC	488
OY	800	AGTATGGGTCCGCGACACTGACGTCCTCA	CCCTGTGTCACTTGCATGATCTCATCTTGAGG	859
Db	489	AGTATGGGTCCGCGACACTGACGTCCTCA	CCCTGTGTCACTTGCATGATCTCATCTTGAGG	548
OY	860	AGGGTGAACCGCTGGAGCCTGGGGAGCTCC	AGCTGAAGAGCTGACTGTGAGAGCCTGAC	919
Db	549	AGGGTGAACCGCTGGAGCCTGGGGAGCTCC	AGCTGAAGAGCTGACTGTGAGAGCCTGAC	608
OY	920	AGCACACGCTGACTCACTGATGACAGATG	AGCTGGCTGTGGCCACCGAGATGGTGTTCAGGC	979
Db	609	AGCACACGCTGACTCACTGATGACAGATG	AGCTGGCTGTGGCCACCGAGATGGTGTTCAGGC	668
OY	980	GGCAGGAATGGTTACGCGACGTGCACAC	AGGAGCTCC-GGAAATGAAGAGAGAGAAC-CCCA	1037
Db	669	GGCAGGAATGGTTACGCGACGTGCACAC	AGGAGCTCCGGGAATGAAGAGAGAGAACCCCCC	728
OY	1038	CCCCCGGAGCGGGTGACCTGCTGGGCA	AGAGGCAAGTGTCTGCACCAAGACACTGACAGG	1097
Db	729	CCCCCGGAGCGGGTGAGCTGCTGGGCA	AGAGGCAAGTGTCTGCACCAAGAGGCTGCACAGG	788
OY	1098	GCTGCAAGTACCGCTGTGACAGCAGAGCC	CAAGCTGACAGGCC--AGCAGAGATGCTGACG	1155
Db	789	GCTGCAAGTACCGCTGTGACAGCAGAGCC	CAAGCTGACAGGCCCAACCAAGAAATGCTGCA	848
OY	1156	ACCAAGCT-GGAGCACTGGGGCCCCGGG	CGAGCCCGGCTGTG----CTGCTCTGCACAG	1210
Db	849	AAACAGCTGGAGACACCGGGGGCCCCGG	GGAGCCCCCCCCCTGGGGTGTGCTCCGGCAGG	908
OY	1211	ATTGACCGGCACTC	1223	
Db	909	AAGAAACCCCCAC	921	

RESULT	12
BMS43870	
LOCUS	
DEFINITION	BMS43870 1114 bp mRNA linear EST-20-FEB-2002
ACCESSION	AGNCNCURT_6492411 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5589213
VERSION	5, mRNA sequence.
KEYWORDS	BMS43870 BMS43870.1 GI:18774639
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 1114) NIH-MGC http://mgc.ncl.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Invitrogen cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be

```

found through the I.M.A.G.E Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL2361 row: 1 column: 22
High quality sequence stop: 566.
Location/Qualifiers
1..1114

```

/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6; Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool of three ovaries, from females ranging in age from 38 to 49 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 2.1 kb, insert size range 1.3-5 kb. Library was normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 036."

ORIGIN

Query Match	31.5%	Score 710.6	DB 4	Length 1114
Best Local Similarity	95.7%	Pred. 31.6e-139		
Matches 794	Conservative 0	Mismatches 27	Indels 9	Gaps 6

OY	147	CACATGTCCTCCGAGGACAGGTGGGGGCGAGCGGGCCATCAAGCCCTGACAGCCCAT	206
Db	12	CCACATGTCCTCGAGGATAGTGGGGGCGAGCGGGCCATCAAGCCCTGACAGCCCAT	71
OY	207	CAGTCAGTCTCTGGGCTGAGATCAACCAAGCAAACTGAAGGGCTGAGCCGCTTGCGGCA	266
Db	72	CAGTCAGTCTCTGGGCTGAGATCAACCAAGCAAACTGAAGGGCTGAGCCGCTTGCGGCA	131
OY	267	GCAGCGAAGGATTTGAACTTCAGGGGCCCCCTGAGCAAGCTGAGCTGCTCATCCGGGAAACG	326
Db	132	GCAGCGAAGGATTTGAACTTCAGGGGCCCCCTGAGCAAGCTGAGCTGCTCATCCGGGAAACG	191
OY	327	GCAGCAGCTTCGCAAGACCTACAGCGAGCAAGTGCCAGCAGCTGCGACAGAGAGCTCACAA	386
Db	192	GCAGCAGCTTCGCAAGACCTACAGCGAGCAAGTGCCAGCAGCTGCGACAGAGAGCTCACAA	251
OY	387	GACCCACAGCCAGGACATTTGAAAGCTGAAAGAGCCAGTAACGAGCTTGCAACGGGACAG	446
Db	252	GACCCACAGCCAGGACATTTGAAAGCTGAAAGAGCCAGTAACGAGCTTGCAACGGGACAG	311
OY	447	TGCCCAAGCCAGCGCAAGTACCAAGAGGCGCAGCAAGACAAAGACCTGTACAAAGGCCAA	506
Db	312	TGCCCAAGCCAGCGCAAGTACCAAGAGGCGCAGCAAGACAAAGACCTGTACAAAGGCCAA	371
OY	507	GGACAATATATGTGGGCGAGCCGTGGGAAAGCTTTTGCTCACCAAAACCGATATGTGCTGGG	566
Db	372	GGACAATATATGTGGGCGAGCCGTGGGAAAGCTTTTGCTCACCAAAACCGATATGTGCTGGG	431
OY	567	CGTGCGGGCTTGCGAGGCTTACCAACCAAGCAACCAACCACTCTGTCGCGGCTGCT	626
Db	432	CGTGCGGGCTTGCGAGGCTTACCAACCAAGCAACCAACCACTCTGTCGCGGCTGCT	491
OY	627	GCGGTCATCTGAGGACCTTGCAACGAGGAGATGGCTTGATCTTGAAGGATCTCTGCAGGA	686
Db	492	GCGGTCATCTGAGGACCTTGCAACGAGGAGATGGCTTGATCTTGAAGGATCTCTGCAGGA	551
OY	687	ATTACTGAGATTATGACAGCCTGGTGCAGATATGAGTGGTGGCAATTACCGGGGAGATGGC	746
Db	552	ATTACTGAGATTATGACAGCCTGGTGCAGATATGAGTGGTGGCAATTACCGGGGAGATGGC	611
OY	747	TGCAGCTGCTGCCCGCATCCAGCCTTGAGGCTGTGATCAAGAGGCTTCTTGCGACAGTATGG	806
Db	612	TGCAGCTGCTGCCCGCATCCAGCCTTGAGGCTGTGATCAAGAGGCTTCTTGCGACAGTATGG	671
OY	807	GTCGCACTTGAGGTCCTCAACCTTGCTGTCAACGTTGATGAAGTCACTGC--TTGAGAGGGT	864

Db 672 GTCCGACCTGACGCTCCACCTGTGTGTCATGATGAACCACTGCTTGAAGGAGATT 731
Qy 865 GAACCCCTGAGAGCT--GGGAGGCTCCAG--TGAAGAGCTGACTGTG--AGAGCGTGCA 920
Db 732 AACCCGCTGAGAGCTGGGAGAGCTCCAGCTTGAACCACTGACTGTGGAAGGCGGCA 791
Qy 921 GCACAGC--TGACCTCAGTGAAGATGAGC--TGGCTGTGCGCACGAGA 967
Db 792 CCACAGCCCTGACCTTATGACCAATGAGCTTGGCTGTGGGACCGAAA 841

RESULT 13

BU184736 847 bp mRNA linear EST 04-SEP-2002
LOCUS BU184736
DEFINITION AGENCOURT_8071838 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6088554
5', mRNA sequence.

ACCESSION BU184736
VERSION BU184736.1 GI:22698720
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 847)
NIH-MGC <http://mgs.nci.nih.gov/>.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb-rc@mail.nih.gov
Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LICM2326 row: 0 column: 19
High quality sequence stop: 721.
Location/Qualifiers

FEATURES

source

1..847
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6088554"
/issue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_112"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA Synthesis Kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN

Query Match 31.2%; Score 703.2; DB 5; Length 847;
Best Local Similarity 99.4%; Pred. No. 1.3e-137;
Matches 716; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 380 TCACCAAGACCCAGCAGCAGCAGTGAAGCTGAAGACCAATACCGAGCTGTGGCAC 439
Db 69 TCACCTCAACCCACAGCAGCAGCAGTGAAGCTGAAGACCAATACCGAGCTGTGGCAC 128
Qy 440 GGGAGAGTGGCCAGAGCCAGGAGCAAGTACAGAGGCGCACAAAGACAGACCGTGACA 499
Db 129 GGGAGAGTGGCCAGAGCCAGGAGCAAGTACAGAGGCGCACAAAGACAGACCGTGACA 188
Qy 500 AGGCGAAGACAGTATGTGGCGCAGCTGTGGAAAGCTTTTGTCTACACCAACCGCTATG 559
Db 189 AGGCGAAGACAGTATGTGGCGCAGCTGTGGAAAGCTTTTGTCTACACCAACCGCTATG 248

Qy 560 TCTGTGGCTGTGGGCTGTGGAGCTGACAGTACACACACAGACCAACAGCTCTGTGCCCC 619
Db 249 TCTGTGGGCTGTGGGCTGTGGAGCTGACAGTACACACACAGACCAACAGCTCTGTGCCCC 308
Qy 620 GCTGTGTGGTCTGCTGCTGACAGACCTGACAGAGAGATGCTTGCATCTCGAAGAGATCC 679
Db 309 GCTGTGTGGTCTGCTGCTGACAGACCTGACAGAGAGATGCTTGCATCTCGAAGAGATCC 368
Qy 680 TGCAGGAATACCTGGAGATTAACAGCTGTGACAGGATGAGGTGTGGCATTTACCCGG 739
Db 369 TGCAGGAATACCTGGAGATTAACAGCTGTGACAGGATGAGGTGTGGCATTTACCCGG 428
Qy 740 AGATGCTGTGACAGCTGTGCTGCCCATTCACAGCTGAGGCTGAGTACCAAGGCTTCTGCGAC 799
Db 429 AGATGCTGTGACAGCTGTGCTGCCCATTCACAGCTGAGGCTGAGTACCAAGGCTTCTGCGAC 488
Qy 800 AGTATGGTGTCCGACCTGACAGTCCCACTGTGTCAAGTTCATGATGATCACTGCTTGAAG 859
Db 489 AGTATGGTGTCCGACCTGACAGTCCCACTGTGTCAAGTTCATGATGATCACTGCTTGAAG 548
Qy 860 AGGCTGAACCGCTGGAGCCTGGGGAGCTCCAGCTGAACGAGCTGATGTGGAGGCGTGC 919
Db 549 AGGCTGAACCGCTGGAGCCTGGGGAGCTCCAGCTGAACGAGCTGATGTGGAGGCGTGC 608
Qy 920 AGCACACGCTGACCTCAGTGAAGATGAGCTGCTGTGCGCACCGAGATGTGTTGAGGC 979
Db 609 AGCACACGCTGACCTCAGTGAAGATGAGCTGCTGTGCGCACCGAGATGTGTTGAGGC 668
Qy 980 GGCAGAGATGTGTACGACAGCTGTGCAACAGAGAGCTCCGGAATGAAGAGAAACACCCACC 1039
Db 669 GGCAGAGATGTGTACGACAGCTGTGCAACAGAGAGCTCCGGAATGAAGAGAAACACCCACC 728
Qy 1040 CCC-GGGAGCGGGTGTGAGCTGTGGGCAAGAGCAAGTCTGCAAGAACACTGTGAGGG 1098
Db 729 CCGGGAGCGGGTGTGAGCTGTGGGCAAGAGCAAGTCTGCAAGAACACTGTGAGGG 788

RESULT 14

BX359009/c

LOCUS BX359009 989 bp mRNA linear EST 08-APR-2004
DEFINITION BX359009 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI051YO20 3-PRIME, mRNA sequence.

ACCESSION BX359009
VERSION BX359009.2 GI:46305682
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 989)
LI, W.B., Gruber, C., Jessee, J., and Polayes, D.

Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30376296.
COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7663.f

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?c=CS0DI051BHI0NP1c=7663.f>.

FEATURES

source

1..989
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI051YO20"
/issue_type="PLACENTA COT 25-NORMALIZED"


```
Qy 1098 GCTGCAAGTAGCGCTGTGTCAGCCAGGCCAAGCTGCAGGCCACAGAGAGTTGCTGCAGAC 1157
Db 481 GCTGCAAGTAGCGCTGTGTCAGCCAGGCCAAGCTGCAGGCCACAGAGAGTTGCTGCAGAC 540
Qy 1158 CAAGCTGAGACACTGGGCCCCGGCGAGAGCCCCCTGTGCTGCTCTGCGAGATGACCG 1217
Db 541 CAAGCTGAGACACTGGGCCCCGGCGAGAGCCCCCTGTGCTGCTCTGCGAGATGACCG 600
Qy 1218 CCACTCCACGTCTGCTCCGAGAGCAAGAGCGAGAGGG-GGGAAGACACCCACGCTGGAGA 1276
Db 601 CCACTCCACGTCTGCTCCGAGAGCAAGAGCGAGAGGAGAGACCCACGCTGGAGA 660
Qy 1277 TCCTTAAGAGCCAATCTCAGAGATCTTCCGCCCAAGTTCTCG 1320
Db 661 TCCTTAAGAGCCAATCTCAGAGATCTTCCGCCCAAGTTCTCG 704
```

Search completed: March 19, 2005, 14:10:46
Job time : 6521.86 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 01:20:31 ; Search time 1076.29 Seconds
(without alignments)
12408.301 Million cell updates/sec

Title: US-10-660-763-1_COPY_72_2327

Perfect score: 2256

Sequence: 1 atggcgtctctctctgcgact.....gcattccgaagcgcgcacg 2256

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq.16dec04:*
1: geneeqn19808:*
2: geneeqn19808:*
3: geneeqn20008:*
4: geneeqn2001as:*
5: geneeqn2001as:*
6: geneeqn2002as:*
7: geneeqn2002as:*
8: geneeqn2003as:*
9: geneeqn2003as:*
10: geneeqn2003as:*
11: geneeqn2003as:*
12: geneeqn2004as:*
13: geneeqn2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2256	100.0	2674	6	ABA94500	Abv94500 Human pro
2	2256	100.0	2674	8	ABV75413	Abv75413 Human kin
3	2256	100.0	2674	12	ADM28580	Adm28580 Human pro
4	2206.2	97.8	2599	12	ADK71888	Adk71888 Human kin
5	2023.4	89.7	2889	6	AA895001	AA895001 Human DNA
6	1671.6	74.1	2623	13	ADR39816	Adr39816 Human kin
7	1540.4	68.3	2680	12	ADH43089	Adh43089 CRAM prot
8	635	28.1	856	1	AAAT70060	AaAT70060 fcs/fps p
9	375	16.6	2818	11	ADL71052	AdL71052 Gene enco
10	348.8	15.5	527	13	ADQ54298	AdQ54298 Novel can
11	328.8	14.6	2584	11	ADL71054	AdL71054 Gene enco
12	326.8	14.5	449	9	ACH14291	Ach14291 Human adu
13	296.2	13.1	361	6	AA817052	Aa817052 Human pro
14	296.2	13.1	361	10	AA61604	Aa61604 Human pro
15	296.2	13.1	361	10	ABX14576	Abx14576 Human pro
16	271.6	12.0	1779	8	ACC57665	Acc57665 Drosophi1
17	271.6	12.0	3198	4	ABL18403	AbL18403 Drosophi1
18	271.6	12.0	3369	4	ABL14457	AbL14457 Drosophi1
19	227.2	10.1	15297	6	ABA94501	Abv94501 Human pro
20	227.2	10.1	15297	8	ABV75414	Abv75414 Human kin

21	227.2	10.1	15297	12	ADM28582	Adm28582 Human pro
22	225.6	10.0	12263	6	ABK84514	Abk84514 Human CDN
23	201.2	8.9	3286	12	ADK71885	Adk71885 Human kin
24	201.2	8.9	3875	6	ABZ35414	Abz35414 Human gen
25	199.6	8.8	5616	10	AD608824	Ad608824 Novel DNA
26	199.2	8.8	2778	12	ADP28218	AdP28218 Human sec
27	198	8.8	2955	6	ABV78117	Abv78117 Human eph
28	198	8.8	2955	6	ABZ35693	Abz35693 Human eph
29	198	8.8	2955	6	ABX09936	Abx09936 Human eph
30	198	8.8	2955	6	ABL91658	AbL91658 Human pol
31	198	8.8	3370	6	ABL67952	AbL67952 Ovary can
32	198	8.8	3370	8	ABZ34861	AbZ34861 Coding se
33	198	8.8	3370	10	ADD18447	Add18447 Human pro
34	198	8.8	3370	10	AD638440	Ad638440 Human pro
35	198	8.8	3370	12	ADJ75094	Adj75094 Marker ge
36	198	8.8	3370	13	ADQ80244	AdQ80244 Eph1 cDNA
37	185.4	8.2	3042	6	ABV78118	Abv78118 Human eph
38	185.4	8.2	3042	6	ABZ35694	Abz35694 Human eph
39	185.4	8.2	3042	6	ABX09937	Abx09937 Human eph
40	185.4	8.2	3042	6	ABX91659	AbX91659 Human pol
41	185.4	8.2	3346	13	ACN42099	Acn42099 Human dia
42	185.4	8.2	3319	13	ACN37278	Acn37278 Tumour-as
43	185.4	8.2	3921	6	ABV94231	Abv94231 Breast ca
44	185.4	8.2	3921	8	ABZ34865	Abz34865 Coding se
45	185.4	8.2	3921	10	ADB37439	Adb37439 Human can

ALIGNMENTS

RESULT 1	ABV94500	ABV94500 standard; cDNA, 2674 BP.
ID	ABV94500	
XX	ABV94500;	
AC	09-APR-2002	(first entry)
DT	09-APR-2002	(first entry)
XX	Human proto-oncogene tyrosine kinase encoding cDNA.	
DE	Human proto-oncogene tyrosine kinase encoding cDNA.	
XX	Proto-oncogene tyrosine kinase; potK; tumour; cytostatic; anti-leukemic;	
KW	gene therapy; protein therapy; vaccine; enzymatic-inhibition; human;	
KW	anti-kinase; ss.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FX	5'UTR	1..71
FT		/*tag= a
FT		72..2330
FT		/*tag= b
FT		/product= "proto-oncogene tyrosine kinase"
FT	3'UTR	2331..2674
FT		/*tag= c
XX	US6340584-B1.	
XX	22-JAN-2002.	
XX	27-MAR-2001; 2001US-00817180.	
XX	27-MAR-2001; 2001US-00817180.	
XX	(PEKE) PE CORP NY.	
XX	Gan W, Ye J, Di Francesco V, Beasley EM;	
XX	WPI, 2002-138497/18.	
XX	P-PSDB; ABB07354.	
XX	Nucleic acids encoding a proto-oncogene tyrosine kinase, useful for the	
XX	prevention, diagnosis and treatment of e.g. leukemia and lung tumors.	

PS Claim 1, Fig 1A-B; 49P; English.

XX The invention provides isolated nucleic acid sequences encoding a proto-oncogene tyrosine kinase (pork). The pork polynucleotides and protein may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate pork expression, such as lung and kidney tumors, leukemia and stomach adenocarcinoma. pork may be used to treat diseases associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of pork by expressing inactive proteins or to supplement the patient's own production of pork. The encoded pork may be used as an antigen in the production of pork antibodies against pork and in assays to identify modulators of pork expression and activity. The anti-pork antibodies and antagonists may be used to down regulate expression and activity and as diagnostic agents for detecting the presence of pork in samples. The present sequence represents a cDNA encoding the human pork

XX Sequence 2674 BP; 662 A; 767 C; 810 G; 435 T; 0 U; 0 Other;

Query Match 100.0%; Score 2256; DB 6; Length 2674;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGCTTCTTCTTGAGCTGTGACGCCCCCAGAGGCGACGGGGTCTTGACGCAATGACG 60
DB 72 ATGGGCTTCTTCTTGAGCTGTGACGCCCCCAGAGGCGACGGGGTCTTGACGCAATGACG 131
QY 61 GAGGCCGACCTTGTCTTACTGAGGGCAGTAAAGAACTGATGGCCAGCGGGTCAAGAT 120
DB 132 GAGGCCGACCTTGTCTTACTGAGGGCAGTAAAGAACTGATGGCCAGCGGGTCAAGAT 191
QY 121 GACAGGAGATATGACGAGCTCTTACCAATGTCCCTGACAGGACAGTGGGGCCAGAGC 180
DB 192 GACAGGAGATATGACGAGCTCTTACCAATGTCCCTGACAGGACAGTGGGGCCAGAGC 251
QY 181 CGGGCCATCAGCCCTTGACAGCCCCCATCACTCACTGCTGAGATCAACGCAAACT 240
DB 252 CGGGCCATCAGCCCTTGACAGCCCCCATCACTCACTGCTGAGATCAACGCAAACT 311
QY 241 GAGGGCTGAGCGCTTGTGCGGGCAGCAGCAGAGAGATCTGAACCTCAGAGGCCCTGAGC 300
DB 312 GAGGGCTGAGCGCTTGTGCGGGCAGCAGCAGAGAGATCTGAACCTCAGAGGCCCTGAGC 371
QY 301 AAGCTGAGCTGCTCTCCGGGAAACGGCAGAGCTTCCGAAAGACTTACAGCAGCAGTGG 360
DB 372 AAGCTGAGCTGCTCTCCGGGAAACGGCAGAGCTTCCGAAAGACTTACAGCAGCAGTGG 431
QY 361 CAGCAGCTCAGAGAGCTCACCAAGACCCACAGCCAGGACATTGAAAGCTGAAGAC 420
DB 432 CAGCAGCTCAGAGAGCTCACCAAGACCCACAGCCAGGACATTGAAAGCTGAAGAC 491
QY 421 CAGTACCGAGCTTGGCAGCGGACAGTGGCCAAAGCCAAAGCGGAAGTACAGAGGCGCAGC 480
DB 492 CAGTACCGAGCTTGGCAGCGGACAGTGGCCAAAGCCAAAGCGGAAGTACAGAGGCGCAGC 551
QY 481 AAAGACAAGACCGGTGACAAGGCCAAGGACAATATATGTGGCAGCTGTGGAAGCTTT 540
DB 552 AAAGACAAGACCGGTGACAAGGCCAAGGACAATATATGTGGCAGCTGTGGAAGCTTT 611
QY 541 GCTCACCACAAACGCTATGTGCTGGGCGTGGCGGCTGCGCAGCTACCAACCAACGACAC 600
DB 612 GCTCACCACAAACGCTATGTGCTGGGCGTGGCGGCTGCGCAGCTACCAACCAACGACAC 671
QY 601 CAGCAACTCTGTGCGCGGCGCTGCTGCGGTCACTGCAAGACCTTGACAGAGAGATGGCT 660
DB 672 CAGCAACTCTGTGCGCGGCGCTGCTGCGGTCACTGCAAGACCTTGACAGAGAGATGGCT 731
QY 661 TGCATCTGAAGAGATCTCTGACAGAAATCTGAGATTAGCAGCTGTGTCAGATGAG 720
DB 732 TGCATCTGAAGAGATCTCTGACAGAAATCTGAGATTAGCAGCTGTGTCAGATGAG 791
QY 721 GTGTGAGCAATTCAGCGGAGATGAGTGCAGCTGTCCCGCATTCAGGCTGAGGCTGAG 780
```

```
DB 792 GTGTGAGCAATTCAGCGGAGATGAGTGCAGCTGTCCCGCATTCAGGCTGAGGCTGAG 851
QY 781 TACCAAGGCTTCTGCGACAGTATGGGTCGCGACCTGACGTCCACCCCTGTGTCAGGTT 840
DB 852 TACCAAGGCTTCTGCGACAGTATGGGTCGCGACCTGACGTCCACCCCTGTGTCAGGTT 911
QY 841 GATGAGTCACTGCTTGAAGAGGGTGAACCGCTGAGACCTGGGAGGCTCCAGCTGAACGAG 900
DB 912 GATGAGTCACTGCTTGAAGAGGGTGAACCGCTGAGACCTGGGAGGCTCCAGCTGAACGAG 971
QY 901 CTGACTGTGAGAGAGGTGACAGCACCCTGACCTTCACTGACAGTATGAGTGTGCTGAGCC 960
DB 972 CTGACTGTGAGAGAGGTGACAGCACCCTGACCTTCACTGACAGTATGAGTGTGCTGAGCC 1031
QY 961 ACCGAGATGAGTTCAGGCGGAGGAGATGTTTACAGCTGCAACAGGAGCTCCGGAAT 1020
DB 1032 ACCGAGATGAGTTCAGGCGGAGGAGATGTTTACAGCTGCAACAGGAGCTCCGGAAT 1091
QY 1021 GAAGAGAGAACACCCACCCCGGAGCGGAGTGCAGCTGTGGCGAAGAGCAATGCTG 1080
DB 1092 GAAGAGAGAACACCCACCCCGGAGCGGAGTGCAGCTGTGGCGAAGAGCAATGCTG 1151
QY 1081 CAAGAGCACTGACAGGGCTGACAGTATGGCTGTGACAGCCAGCCAAAGCTGACAGCCGAG 1140
DB 1152 CAAGAGCACTGACAGGGCTGACAGTATGGCTGTGACAGCCAAAGCTGACAGCCGAG 1211
QY 1141 CAGAGTTGCTGACAGCAACCAAGCTGAGAGCACTGGGCGCGGAGGCCCCGCTGTGCTG 1200
DB 1212 CAGAGTTGCTGACAGCAACCAAGCTGAGAGCACTGGGCGCGGAGGCCCCGCTGTGCTG 1271
QY 1201 CTCTGACAGATGACCGCCACTTCACTGCTCTGGAGCAGAGAGCAGAGGGGGAGAG 1260
DB 1272 CTCTGACAGATGACCGCCACTTCACTGCTCTGGAGCAGAGAGCAGAGGGGGAGAG 1331
QY 1261 ACACCACGCTGAGATCTTAAAGCCCACTTCAAGAAATCTTCCGCCCAAGTCTG 1320
DB 1332 ACACCACGCTGAGATCTTAAAGCCCACTTCAAGAAATCTTCCGCCCAAGTCTG 1391
QY 1321 AACCTGTACCGACTGGAAGGGAGAGGCTTCTAGCACTCTTGTCTCATGCAACCACTA 1380
DB 1392 AACCTGTACCGACTGGAAGGGAGAGGCTTCTAGCACTCTTGTCTCATGCAACCACTA 1451
QY 1381 CTGAGCACCCAGCAGCCCTTCAACAAAGAGTGTGTCTGTGACAGGGCTGTGCC 1440
DB 1452 CTGAGCACCCAGCAGCCCTTCAACAAAGAGTGTGTGTCTGTGACAGGGCTGTGCC 1511
QY 1441 AAGACAAGTGGGTCTGAACCATGAGGACCTGTGTGTGGTGAAGCAATTGACAGGGG 1500
DB 1512 AAGACAAGTGGGTCTGAACCATGAGGACCTGTGTGTGGTGAAGCAATTGACAGGGG 1571
QY 1501 AACTTGGGGAAGTGTTCAGCGGACGCTGCGAGCCGACCAACCTGTGGCGGTGAAG 1560
DB 1572 AACTTGGGGAAGTGTTCAGCGGACGCTGCGAGCCGACCAACCTGTGGCGGTGAAG 1631
QY 1561 TCTTGTGAGAGACGCTCCACCTGACTCAAGAGCCCAAGTCTTCAAGAAAGCGAGATC 1620
DB 1632 TCTTGTGAGAGACGCTCCACCTGACTCAAGAGCCCAAGTCTTCAAGAAAGCGAGATC 1691
QY 1621 CTGAAGCAATACAGCACCCCAACATCTGTGCGCTTCAATTGTGTCTGCAACCCAGAGAG 1680
DB 1692 CTGAAGCAATACAGCACCCCAACATCTGTGCGCTTCAATTGTGTCTGCAACCCAGAGAG 1751
QY 1681 CCATCTACATGTCATGAGAGCTTGTGACGGGGGGGAGATTCTCTGACCTTCTCGGACG 1740
DB 1752 CCATCTACATGTCATGAGAGCTTGTGACGGGGGGGAGATTCTCTGACCTTCTCGGACG 1811
QY 1741 GAGGGGCGCGCTGCGGGTGAAGACTGTGCTGACAGATGTGGGGATGCAAGTGTGCGC 1800
DB 1812 GAGGGGCGCGCTGCGGGTGAAGACTGTGCTGACAGATGTGGGGATGCAAGTGTGCGC 1871
QY 1801 ATGAGATCTGGAAGAGCAAGTGTGATCAACCGGAGACCTGTGCTCGGAATGCTG 1860
DB 1872 ATGAGATCTGGAAGAGCAAGTGTGATCAACCGGAGACCTGTGCTCGGAATGCTG 1931
```

Qy	1861	GTGACAGAGAAAGATGTCTCTGAAATCATAGTACTTTGGGAGTGTCCCGAGAGAAAGCCGAT	1920
Dp	1932	GTGACAGAGAAAGATGTCTCTGAAATCATAGTACTTTGGGAGTGTCCCGAGAGAAAGCCGAT	1991
Qy	1921	GGGGCTCATGCAAGCCTCAGGGGGGCTCAGACMAGTCCCGTGAAGTGAGACCGACCTGAG	1980
Dp	1992	GGGGTCTATGACAGCCTCAGGGGGGCTCAGACMAGTCCCGTGAAGTGAGACCGACCTGAG	2051
Qy	1981	GCCTTTAACTACGGGCGCTACTCTCTCGAAAGCGACGTGTGAGCTTTGGCATCTTGCTC	2040
Dp	2052	GCCCTTAACTACGGGCGCTACTCTCTCGAAAGCGACGTGTGAGCTTTGGCATCTTGCTC	2111
Qy	2041	TGGAGAGACTTCAGCCTGGGGGGCTCCCGCTATCCAACTCAGCAATCAGACAGACAGG	2100
Dp	2112	TGGAGAGACTTCAGCCTGGGGGGCTCCCGCTATCCAACTCAGCAATCAGACAGACAGG	2171
Qy	2101	GAGTTTGTGAGAAAGGGGGGGCGGTCTGCGCTGCCAGAGCTGTCTGATGCGGTGTC	2160
Dp	2172	GAGTTTGTGAGAAAGGGGGGGCGGTCTGCGCTGCCAGAGCTGTGTCTGATGCGGTGTC	2231
Qy	2161	AGGCTCATGAGAGCATGTGTGGGCGCTATGAGCGCTGGGACAGCGGCCCATGTTTACAGCACATC	2220
Dp	2232	AGGCTCATGAGAGCATGTGTGGGCGCTATGAGCGCTGGGACAGCGGCCCATGTTTACAGCACATC	2291
Qy	2221	TACACAGAGCTGCAGAGCATCCGAAAGCGGCGATCGG	2256
Dp	2292	TACACAGAGCTGCAGAGCATCCGAAAGCGGCGATCGG	2327

RESULT 2
ABV75413
ID ABV75413 standard; cDNA; 2674 BP

24-JAN-2003	(first entry)
Human kinase encoding cDNA.	
Human; kinase; proto-oncogene tyrosine kinase; antiarteriosclerotic; cytostatic; antiinflammatory; antipsoriatic; gene therapy;	
protein kinase; drug screening assay; tissue typing; chromosome 15;	
pharmacogenomic analysis; inflammation; cancer; leukaemia; lung tumours;	
kidney tumour; stomach adenocarcinoma; arteriosclerosis; psoriasis; gene	
ss.	
Homo sapiens.	
Key	Location/Qualifiers
5'UTR	1..71
FT	/*tag= a
CDS	72..2330
FT	/*tag= b
FT	/product= "protein kinase"
3'UTR	2331..2674
FT	/*tag= c
MO20027191-A2.	
03-OCT-2002.	
27-MAR-2002; 2002MO-US009325.	
27-MAR-2001; 2001US-00817180.	
06-DEC-2001; 2001US-00003295.	
(PEKE) PE CORP NY.	
Gan W, Ye J, Di Francesco V, Beasley EM;	
WPI, 2003-029927/02.	
P-PsDB; ABB99046.	

New human kinase protein, useful for treating or diagnosing disorders associated with an absence of, inappropriate, or unwanted expression of the protein, e.g., inflammation or cancer, in drug screening assays and pharmacogenomics.

Claim 4 (b) ; Fig 1, 75pp; English.

The invention relates to a newly isolated peptide sequence of a human kinase that is related to the proto-oncogene tyrosine kinase subfamily. The activity of the kinase of the invention may be described as, cytosolic, antiarteriosclerotic, antiinflammatory and antiporiatic. Peptides of the invention are useful in assays to determine the biological activity of the protein, in drug screening assays, tissue typing and pharmacogenomic analysis. They are also useful in treating or diagnosing disorders characterised by an absence of, inappropriate, or unadapted expression of the protein, such as inflammation, cancer (e.g. leukaemia, lung tumours, kidney tumours or stomach adenocarcinoma), arteriosclerosis, and psoriasis. Nucleic acid molecules of the invention are useful as probes, primers and chemical intermediates in biological assays. The peptide and nucleic acid sequences are useful as models for the development of human therapeutic targets, aid in the identification of therapeutic proteins and serve as targets for the development of human therapeutic agents that modulate kinase protein activity in cells and tissues that express the protein. The protein of the invention may also be useful in gene therapy. The gene encoding the protein of the invention has been localised to human chromosome 15. The current sequence represents the human kinase encoding cDNA

SQ Sequence 2674 BP; 662 A; 767 C; 810 G; 435 T; 0 U; 0 Other;

Query Match	100.0%;	Score 2256;	DB 8;	Length 2674;
-------------	---------	-------------	-------	--------------

Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGGCTTCTCTTTCGAGCTGTGCAGACCCGACGGGCGACACGGGGTCTCTCAGCAAAATGAG	60
Db	72	ATGGGCTTCTCTTTCGAGCTGTGCAGACCCGACGGGCGACACGGGGTCTCTCAGCAAAATGAG	133
QY	61	GAGCGCGAGCTTCTGTCTATCTGGAGGGCATGAGAAAGTGAATGGCCAGCGGCTCAGAGCT	122
Db	132	GAGCGCGAGCTTCTGTCTATCTGGAGGGCATGAGAAAGTGAATGGCCAGCGGCTCAGAGCT	191
QY	121	GACAGGAGATATGACAGGACTGCTTACACCATATGCCCTGACAGGACAGTGGGGGGCAGAGC	180
Db	192	GACAGGAGATATGACAGGACTGCTTACACCATATGCCCTGACAGGACAGTGGGGGGCAGAGC	253
QY	181	CGGGCCATCAGGCGCTTGACAGACCCCATCAGTCAGTCTGGGCTGAGATCACCAGCCAACT	244
Db	252	CGGGCCATCAGGCGCTTGACAGACCCCATCAGTCAGTCTGGGCTGAGATCACCAGCCAACT	311
QY	241	GAGGGCTTGAGCCGCTTGTCTGCTCGGCAGCAGCCGACAGATCTGAACTCAGGGCCCTTGAGC	300
Db	312	GAGGGCTTGAGCCGCTTGTCTGCTCGGCAGCAGCCGACAGATCTGAACTCAGGGCCCTTGAGC	373
QY	301	AAGCTGAGCGTGTCTCATCCGGGAGCGGCGAGCGTCTCCAAAGACCTTACAGCGAGCAGTGG	360
Db	372	AAGCTGAGCGTGTCTCATCCGGGAGCGGCGAGCGTCTCCAAAGACCTTACAGCGAGCAGTGG	433
QY	361	CAGCAGCTSCAGCAGGAGCTCAACCAAGCCACAGCCGACGACATTGAGAGCTGAAAGGC	422
Db	432	CAGCAGCTSCAGCAGGAGCTCAACCAAGCCACAGCCGAGCATTGAGAACTGAAAGGC	493
QY	421	CAGTACCGAGCTCTGGCAGCGGACAGTCCCAAGCCCAAGCCGCAAGTATCAGAGGGCCAGC	480
Db	492	CAGTACCGAGCTCTGGCAGCGGACAGTCCCAAGCCCAAGCCGCAAGTATCAGAGGGCCAGC	553
QY	481	AAAGACAAGACCGTGCACAAAGCCCAAGACAAGTATGTGCAGCGCTGTGGAAGCTTTT	544
Db	552	AAAGACAAGACCGTGCACAAAGCCCAAGACAAGTATGTGCAGCGCTGTGGAAGCTTTT	611
QY	541	GCTCACCAACAACCGCTATGTGCTGGGCGTGGGGCTGGCAGCTACACACCAACACAC	600
Db		GCTCACCAACAACCGCTATGTGCTGGGCGTGGGGCTGGCAGCTACACACCAACACAC	

Db 612 GCTACCAACAACCGATGATGCTGAGGCGTGGCGGCTGCGACAGTACACCAACCAACAC 671
 Qy 601 GACCAAGCTCTGCTGCGCGCGCTGCTGCGGTCACTGCAAGAACTTGACAAGAGAGATGGT 660
 Db 672 CACCAAGCTCTGCTGCGCGCGCTGCTGCGGTCACTGCAAGAACTTGACAAGAGAGATGGT 731
 Qy 661 TGCATCTGAAGAGAGATCTGCAAGAAATCCTGGAATTAAGACGCTGGTGCAAGATGAG 720
 Db 732 TGCATCTGAAGAGAGATCTGCAAGAAATCCTGGAATTAAGACGCTGGTGCAAGATGAG 791
 Qy 721 GTGTGGCCATTCACCGGAGAGATGCTGCAAGTCTGCTGCGCATCCAGCTTGAAGCTGAG 780
 Db 792 GTGTGGCCATTCACCGGAGAGATGCTGCAAGTCTGCTGCGCATCCAGCTTGAAGCTGAG 851
 Qy 781 TACCAAGGCTTCCGTGGAAGATGAGTGGTCCGACCTGAGAGTCCACCTGTGTCAAGTTC 840
 Db 852 TACCAAGGCTTCCGTGGAAGATGAGTGGTCCGACCTGAGAGTCCACCTGTGTCAAGTTC 911
 Qy 841 GATGAGTCACTGCTTGAAGAGAGGTGAACCGCTGGAAGCTGGGAGCTCGAGCTGAACGAG 900
 Db 912 GATGAGTCACTGCTTGAAGAGAGGTGAACCGCTGGAAGCTGGGAGCTCGAGCTGAACGAG 971
 Qy 901 CTGACTGTGGAAGAGCTGGAAGACACGCTGACCTCAAGTGAAGATGAGCTGGCTGTGGCC 960
 Db 972 CTGACTGTGGAAGAGCTGGAAGACACGCTGACCTCAAGTGAAGATGAGCTGGCTGTGGCC 1031
 Qy 961 ACCGAGATGCTGTTCAGGCGGCGAGAGATGTTAACGACCTGGAAGAGAGCTCCGGAAAT 1020
 Db 1032 ACCGAGATGCTGTTCAGGCGGCGAGAGATGTTAACGACCTGGAAGAGAGCTCCGGAAAT 1091
 Qy 1021 GAAGAGAGAAACACCAACCCCGGAGCGGAGTCACTGCTGGGCAAGAGGCAAGTGTCTG 1080
 Db 1092 GAAGAGAGAAACACCAACCCCGGAGCGGAGTCACTGCTGGGCAAGAGGCAAGTGTCTG 1151
 Qy 1081 CAAGAAGCACTGAGAGGCGCTGCAAGTGAAGCTGTGCAAGCCAGGCTGCAAGCTCCAG 1140
 Db 1152 CAAGAAGCACTGAGAGGCGCTGCAAGTGAAGCTGTGCAAGCCAGGCTGCAAGCTCCAG 1211
 Qy 1141 CAGGAGTTGCTGAGACCAAGCTGGAAGCACTGGGCGCGGCGAGAGCCCGCTGTGCTG 1200
 Db 1212 CAGGAGTTGCTGAGACCAAGCTGGAAGCACTGGGCGCGGCGAGAGCCCGCTGTGCTG 1271
 Qy 1201 CTCCTGCAAGATGACCGCACTCCAGCTGCTCTCGAGAGAGAGAGAGGAGGAGAG 1260
 Db 1272 CTCCTGCAAGATGACCGCACTCCAGCTGCTCTCGAGAGAGAGAGAGGAGGAGAG 1331
 Qy 1261 ACAACCAAGCTGGAATCTTAAAGACCAATCTCAAGAACTTCCGCCCAAGTTCTCG 1320
 Db 1332 ACAACCAAGCTGGAATCTTAAAGACCAATCTCAAGAACTTCCGCCCAAGTTCTCG 1391
 Qy 1321 AACCTGTACCGACTGGAAGGAGGCTTCCAGATTCCTTGTCTCATCCAGCACTTA 1380
 Db 1392 AACCTGTACCGACTGGAAGGAGGCTTCCAGATTCCTTGTCTCATCCAGCACTTA 1451
 Qy 1381 CTGAGACACCAAGAGCCCTCACCAGAAAGATGCTGTCTGCAAGAGGCTGTGCC 1440
 Db 1452 CTGAGACACCAAGAGCCCTCACCAGAAAGATGCTGTCTGCAAGAGGCTGTGCC 1511
 Qy 1441 AAGGACAAGTGGTCTGAACCATGAGAGACTGTGTGGTGAGCAATTTGACCGAGG 1500
 Db 1512 AAGGACAAGTGGTCTGAACCATGAGAGACTGTGTGGTGAGCAATTTGACCGAGG 1571
 Qy 1501 AACTTTGGGGAATGTTCAAGCGGAGCGCTGCGAGCGCAACACCTGTGGGCGGTGAAG 1560
 Db 1572 AACTTTGGGGAATGTTCAAGCGGAGCGCTGCGAGCGCAACACCTGTGGGCGGTGAAG 1631
 Qy 1561 TCTTGTGAGAGAGCTCCCACTGACCTCAAGGCAAGTTTCTACAGAAAGGAGATC 1620
 Db 1632 TCTTGTGAGAGAGCTCCCACTGACCTCAAGGCAAGTTTCTACAGAAAGGAGATC 1691
 Qy 1621 CTGAGACAGTGAACCAACCCCAACATCTGTGCTCTCATTTGGTGTCTGCAACCAAGACAG 1680
 Db 1692 CTGAGACAGTGAACCAACCCCAACATCTGTGCTCTCATTTGGTGTCTGCAACCAAGACAG 1751

Qy 1681 CCCATCTACATGATGAGAGCTTGGACAGGAGGAGCACTTCTGACCTTCTCGCAG 1740
 Db 1752 CCCATCTACATGATGAGAGCTTGGACAGGAGGAGCACTTCTGACCTTCTCGCAG 1811
 Qy 1741 GAGGAGGCGCGCTCGGAGGAGCACTTCTGACAGATGTTGGAGATGACAGCTGTGC 1800
 Db 1812 GAGGAGGCGCGCTCGGAGGAGCACTTCTGACAGATGTTGGAGATGACAGCTGTGC 1871
 Qy 1801 ATGAGATACCTGGAAGAGCAAGTGTGATTCACCCGAGACCTGCTGTGGAATGCTCC 1860
 Db 1872 ATGAGATACCTGGAAGAGCAAGTGTGATTCACCCGAGACCTGCTGTGGAATGCTCC 1931
 Qy 1861 GTGACAGAGAAATGTCCTGAAAGATCAAGTGAATCTTGGAGTGTCCGAGAGAGCCGAT 1920
 Db 1932 GTGACAGAGAAATGTCCTGAAAGATCAAGTGAATCTTGGAGTGTCCGAGAGAGCCGAT 1991
 Qy 1921 GGGGCTTATGACAGCTCAGAGGAGGCTCAGACAAAGTCCCGTGAAGTGAACCGACCTGAG 1980
 Db 1992 GGGGCTTATGACAGCTCAGAGGAGGCTCAGACAAAGTCCCGTGAAGTGAACCGACCTGAG 2051
 Qy 1981 GCCCTTAATACAGGCGCTAATCTCTCCGAAGCGAGCTGTGAGAGCTTGGACATCTTGTCTC 2040
 Db 2052 GCCCTTAATACAGGCGCTAATCTCTCCGAAGCGAGCTGTGAGAGCTTGGACATCTTGTCTC 2111
 Qy 2041 TGGAGAACCTTCAAGCTGAGGAGGCTCCCGCTAATCCCAACCTGACAGATCAGACAGACG 2100
 Db 2112 TGGAGAACCTTCAAGCTGAGGAGGCTCCCGCTAATCCCAACCTGACAGATCAGACAGACG 2171
 Qy 2101 GAGTTTGTGAGAGAGGAGGCGCTGTGCTGCTCCAGAGCTGTCTGATGCCGTGTC 2160
 Db 2172 GAGTTTGTGAGAGAGGAGGCGCTGTGCTGCTCCAGAGCTGTCTGATGCCGTGTC 2231
 Qy 2161 AAGCTCATGAGACATGCTGTGGGCTATGAGCTGTGGGAGAGGCGGCACTTCAAGACATC 2220
 Db 2232 AAGCTCATGAGACATGCTGTGGGCTATGAGCTGTGGGAGAGGCGGCACTTCAAGACATC 2291
 Qy 2221 TACCAAGAGCTGACAGAGATCCGAAAGCGGATCCG 2256
 Db 2292 TACCAAGAGCTGACAGAGATCCGAAAGCGGATCCG 2327

Db
 RESULT 3
 ADM28580
 ID ADM28580 Brandard; cDNA; 2674 BP.
 XX
 AC ADM28580;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Human proto-oncogene tyrosine kinase cDNA.
 XX
 KW ss; gene; gene therapy; human; proto-oncogene; tyrosine kinase; cancer.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..71
 FT /*tag= a
 FT CDS 72..2330
 FT /*tag= b
 FT /product= "Proto-oncogene tyrosine kinase"
 FT 3'UTR 2331..2674
 FT /*tag= C
 FT
 XX
 PN US2004063130-A1.
 XX
 PD 01-APR-2004.
 XX
 PF 12-SEP-2003; 2003US-00660763.
 XX
 PR 27-MAR-2001; 2001US-00817180.
 PR 06-DEC-2001; 2001US-00003295.

XX (APPL-) APPLERA CORP.
XX Gan W, Ye J, Di Francesco V, Beasley EM,
PI WPI: 2004-282461/26.
XX DR P-PDB; ADM28581.
XX
PT New human kinase peptides, useful for preparing a composition for
PT treating a disease or condition mediated by human kinases.
XX
PS Claim 4; SEQ ID NO 1; 53pp; English.
XX
CC The invention relates to a new isolated human proto-oncogene tyrosine
CC kinase peptide or its allelic variant, orthologue or fragment. The
CC peptide is useful for preparing a composition for treating a disease or
CC condition mediated by a human kinase protein e.g. cancer. The present
CC sequence represents the human proto-oncogene tyrosine kinase cDNA.
XX
SQ Sequence 2674 BP; 662 A; 767 C; 810 G; 435 T; 0 U; 0 Other;
Query Match 100.0%; Score 2256; DB 12; Length 2674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCTTCTCTTCTAGCTGTGACGCCCCAGGCGCCACGGGGTCTCTGACGAAATGACG 60
DB 72 ATGGGCTTCTCTTCTAGCTGTGACGCCCCAGGCGCCACGGGGTCTCTGACGAAATGACG 120
QY 61 GAGGCGGAGCTTCTCTCTAGTGAAGGCGATGAGATGGGCCCGGCTCAAGAT 120
DB 132 GAGGCGGAGCTTCTCTCTAGTGAAGGCGATGAGATGGGCCCGGCTCAAGAT 131
QY 121 GACAGGAGATGACAGAGCTCTTCAACCATGTCTCTGACGACAGTGGGCGCAGAGC 180
DB 192 GACAGGAGATGACAGAGCTCTTCAACCATGTCTCTGACGACAGTGGGCGCAGAGC 251
QY 181 CGGGCCATCAGCCCTGACAGCCCATCATGCTAGTCTCTGGGCTGAGATCACCAAGCAACT 240
DB 252 CGGGCCATCAGCCCTGACAGCCCATCATGCTAGTCTCTGGGCTGAGATCACCAAGCAACT 311
QY 241 GAGGGCTGAGCGCTTGTGCGGCGACGACGAGAGGATCTGAACCTCAGGGCCCTGAGC 300
DB 312 GAGGGCTGAGCGCTTGTGCGGCGACGACGAGAGGATCTGAACCTCAGGGCCCTGAGC 371
QY 301 AAGCTGAGCTGTCTATCCGCGAAGCGGACAGCTTTCGAAAGACTTACACGAGCAGTGG 360
DB 372 AAGCTGAGCTGTCTATCCGCGAAGCGGACAGCTTTCGAAAGACTTACACGAGCAGTGG 431
QY 361 CAGCAGCTGACAGAGAGCTCACCAAGACCCACAGCCAGGACATTGAGAGCTGAAAGAC 420
DB 432 CAGCAGCTGACAGAGAGCTCACCAAGACCCACAGCCAGGACATTGAGAGCTGAAAGAC 491
QY 421 CAGTACCGAGCTGTGCGAGCGGACAGTGCCCAAGGCGAAGTACCGAGGCGCAGC 480
DB 492 CAGTACCGAGCTGTGCGAGCGGACAGTGCCCAAGGCGAAGTACCGAGGCGCAGC 551
QY 481 AAAGACAAGGACCGTGAACAAGCCAAAGACAAATATGTGCGACCTGTGGAAGCTTTT 540
DB 552 AAAGACAAGGACCGTGAACAAGCCAAAGACAAATATGTGCGACCTGTGGAAGCTTTT 611
QY 541 GCTCACCAACAACCGCTATGTGCTGGGGCTGGGGCTGCGAGCTTACCAACCAACAC 600
DB 612 GCTCACCAACAACCGCTATGTGCTGGGGCTGGGGCTGCGAGCTTACCAACCAACAC 671
QY 601 CACCAAGCTTCTGTGCGCGGCTGTGCGGCTCACTGAGGACCTGACGAGAGAGTGGCT 660
DB 672 CACCAAGCTTCTGTGCGCGGCTGTGCGGCTCACTGAGGACCTGACGAGAGAGTGGCT 731
QY 661 TGCATCTGAAGAGATCTCTGAGAAATCTTGAAGATTAGCAGCTGTGTGAGATGAG 720
DB 732 TGCATCTGAAGAGATCTCTGAGAAATCTTGAAGATTAGCAGCTGTGTGAGATGAG 791

QY 721 GTGTGGCCATTACCGGGAGATGGCTGACGCTGTCCCGCATTCACGCTGAGGCTGAG 780
DB 792 GTGTGGCCATTACCGGGAGATGGCTGACGCTGTCCCGCATTCACGCTGAGGCTGAG 851
QY 781 TACCAAGGCTTCTGTGCGAGAGTATGGGTCGACACTGACGTCCACCTGTGTACGTTG 840
DB 852 TACCAAGGCTTCTGTGCGAGAGTATGGGTCGACACTGACGTCCACCTGTGTACGTTG 911
QY 841 GATGAGTCACTGCTTGAAGAGAGGTGAACCGCTGAGCCTGGGGAGCTCCAGTGAACGAG 900
DB 912 GATGAGTCACTGCTTGAAGAGAGGTGAACCGCTGAGCCTGGGGAGCTCCAGTGAACGAG 971
QY 901 CTGACGTGTGAAGGCTGTGACGACACGCTGACCTTCACTGACACATGAGCTGGCTGGCC 960
DB 972 CTGACGTGTGAAGGCTGTGACGACACGCTGACCTTCACTGACACATGAGCTGGCTGGCC 1031
QY 961 ACCGAGATGGTGTTCAGCGCGCAGAGATGGTATACCACTGACGTCGAAACGAGCTCCGGAAT 1020
DB 1032 ACCGAGATGGTGTTCAGCGCGCAGAGATGGTATACCACTGACGTCGAAACGAGCTCCGGAAT 1091
QY 1021 GAAGAGAGAACCCACACCCCGGAGCGGCTGACCTGTGGGCAAGAGCAAGTGTCTG 1080
DB 1092 GAAGAGAGAACCCACACCCCGGAGCGGCTGACCTGTGGGCAAGAGCAAGTGTCTG 1151
QY 1081 CAAAGAGCACTGAGAGGGGCTGACGCTGAGCTGTGACCCAGGCCCAAGCTGACGCGCCAG 1140
DB 1152 CAAAGAGCACTGAGAGGGGCTGACGCTGAGCTGTGACCCAGGCCCAAGCTGACGCGCCAG 1211
QY 1141 CAGGATGCTGTGACAGCAACAACTGAGACACTTGGGCGCGGCGAGCCCGCTGTGCTG 1200
DB 1212 CAGGATGCTGTGACAGCAACAACTGAGACACTTGGGCGCGGCGAGCCCGCTGTGCTG 1271
QY 1201 CTCCTGACAGATGACCGCACTTCACGCTGTCTCTGAGAGGAGCGAGAGGAGGAGG 1260
DB 1272 CTCCTGACAGATGACCGCACTTCACGCTGTCTCTGAGAGGAGCGAGAGGAGGAGG 1331
QY 1261 ACAACCAAGCTGAGATCTCTTAAGACCAATTCAGAGAACTTTCCGCCCAAGTTCTCG 1320
DB 1332 ACAACCAAGCTGAGATCTCTTAAGACCAATTCAGAGAACTTTCCGCCCAAGTTCTCG 1391
QY 1321 AACCTGTAACGATGGAAGGAGGAGGCTTCTTACGATTCCTTGTGCTATCGACCACTTA 1380
DB 1392 AACCTGTAACGATGGAAGGAGGAGGCTTCTTACGATTCCTTGTGCTATCGACCACTTA 1451
QY 1381 CTGAGACCCAGAGGCGCTTCAACCAAGAAAGTGTGTCTCTGCAACAGGCTGTGCCC 1440
DB 1452 CTGAGACCCAGAGGCGCTTCAACCAAGAAAGTGTGTCTCTGCAACAGGCTGTGCCC 1511
QY 1441 AAGGACAGTGGGTGCTGAACCTTGAAGACCTGTGTGTGGGTGAAGATTTGACCGGGG 1500
DB 1512 AAGGACAGTGGGTGCTGAACCTTGAAGACCTGTGTGTGGGTGAAGATTTGACCGGGG 1571
QY 1501 AACTTGGCGAAGTGTTCAGCGGACGCTGCGAGCGGACCAACAACCTGTGTGGGTGAAG 1560
DB 1572 AACTTGGCGAAGTGTTCAGCGGACGCTGCGAGCGGACCAACAACCTGTGTGGGTGAAG 1631
QY 1561 TCTTGTGAGAGAGCGTCCACCTGAACCTTCAAGGCCAAGTTTCTACAGAGGAGATC 1620
DB 1632 TCTTGTGAGAGAGCGTCCACCTGAACCTTCAAGGCCAAGTTTCTACAGAGGAGATC 1691
QY 1621 CTGAAGCAATTAACGACCCCAACATCGTGTGCTTATTTGGTGTGCAACCCAGAAAGCAG 1680
DB 1692 CTGAAGCAATTAACGACCCCAACATCGTGTGCTTATTTGGTGTGCAACCCAGAAAGCAG 1751
QY 1681 CCCATCTACATGCTGTGAGAGCTTGTGAGAGGGGGGAGCTTCTGTACCTTCCCGAAG 1740
DB 1752 CCCATCTACATGCTGTGAGAGCTTGTGAGAGGGGGGAGCTTCTGTACCTTCCCGAAG 1811
QY 1741 GAGGGGGCGCGCTGCGGGGTGAAGACTTGTCTGACAGATGAGTGGGAGATGACGCTGTGGC 1800
DB 1812 GAGGGGGCGCGCTGCGGGGTGAAGACTTGTCTGACAGATGAGTGGGAGATGACGCTGTGGC 1871
QY 1801 ATGGAATTAAGAGAGCAAGTGTGATTCACCGGGACCTGTGCTGTGGAATGCTGCTG 1860

Db	1872	ATGAGATACCTGGAGAGCAAGTGTCTGCATCCACCGGGACTTGGCTGCTCGGAACCTGGCTG	1931
QY	1861	GTGACAGAGAAAGAAATGTCTCTGGAAGATCAGTACTTTGGGATGTCCCGAGAGAAAGCCGAT	1920
Db	1932	GTGACAGAGAAAGAAATGTCTGGAATATCAGTACTTTGGGATGTCCCGAGAGAAAGCCGAT	1991
QY	1921	GGGCTCTATGACGCTTCAGGGGGGCTCAGACAAGTCCCCGTGAAGTGAACCGCACCTTAG	1980
Db	1992	GGGCTCTATGACGCTCAGGGGGGCTCAGACAAGTCCCCGTGAAGTGAACCGCACCTTAG	2051
QY	1981	GCCCTTAACCTACGGGCGGCTACTCTCTCCGAAGCGACGCTGTGACCTTTGGATCTTGCTC	2040
Db	2052	GCCCTTAACCTACGGGCGGCTACTCTCTCCGAAGCGACGCTGTGACCTTTGGATCTTGCTC	2111
QY	2041	TGGAGACCTTCAGGCTGGGGGCGTCCCCCTATCCGAACCTCAGCAATCAGCAGACACGG	2100
Db	2112	TGGAGACCTTCAGGCTGGGGGCGTCCCCCTATCCGAACCTCAGCAATCAGCAGACACGG	2171
QY	2101	GAGTTTGTGAGAAAGGGGGGCGGTCTGGCCCTGCCAGACAGTGTGTCTGATGCGGTGTC	2160
Db	2172	GAGTTTGTGAGAAAGGGGGGCGGTCTGGCCCTGCCAGACAGTGTGTCTGATGCGGTGTC	2231
QY	2161	AGGCTCATGAGACAGTGTGGGCGCTATGAGCCTTGGGACAGCGGCCCAAGCTTGACACATC	2220
Db	2232	AGGCTCATGAGACAGTGTGGGCGCTATGAGCCTTGGGACAGCGGCCCAAGCTTGACACATC	2291
QY	2221	TACCAAGAGCTGCAGAGCATCCGAAGGGGATCGG	2256
Db	2292	TACCAAGAGCTGCAGAGCATCCGAAGGGGATCGG	2327

ID	ADK71888	standard; cDNA; 2599 BP.
XX	ADK71888;	
DT	20-MAY-2004	(first entry)
XX		
DE	Human kinase and phosphatase KPP-6 cDNA.	
XX		
KW	human; kinase; phosphatase; KPP; cardiovascular; antiarteriosclerotic;	
KW	hypotensive; vasotropic; antiinflammatory; antitumoral; anti-HIV;	
KW	antiallergic; antiaesthetic; immunosuppressive; antileukoid;	
KW	dermatologic; antidiabetic; nephroprotective; antigen; gastrointestinal;	
KW	neuroprotective; osteoporotic; antiarthritic; uterine; ophthalmological;	
KW	antitumoral; antiparasitism; nootropic; anticonvulsant; hepatotropic;	
KW	antiparasitic; haemostatic; cytostatic; antipneumatic; antiparasitic;	
KW	antihelminthic; antibacterial; viricide; protozoacide; fungicide;	
KW	cardiovascular disease; immune system; neurological; growth; development;	
KW	cell proliferation; viral; bacterial; fungal; parasitic; protozoan;	
KW	helminthic infection; transgenic; gene therapy; see; gene;	
KW	single nucleotide polymorphism; SNP.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	location/Qualifiers
FT	variation	replace(109,T)
FT		/tag= a
XX		/label= Single nucleotide polymorphism (SNP)
XX		
XX	WO2004018641-A2.	
PD		
XX		
PD	04-MAR-2004.	
XX		
PF	25-AUG-2003;	2003WO-US026635.
XX		
XX	26-AUG-2002;	2002US-0406172P.
PR	25-SEP-2002;	2002US-0413910P.
PR	27-SEP-2002;	2002US-0414296P.
XX		
XX	11-OCT-2002;	2002US-0417821P.
XX		

PA	(INCY-) INCYTE CORP.
XX	
PI	Baughn MR, Richardson TW, Marguis JP, Swarnakar A, Teng YT;
PI	Becha SD, Emerging BM, Jin P, Wilson AD, Yue H, Giesen KU;
PI	Chang H, Yang YG, Lee SY, Khare R, Elliott VS, Hafalia AJA;
PI	Chawla NK, Ramkumar J, Gururajan R, Tribouley CM, Chien D, Tran UK;
PI	Murage J;
XX	
DR	WPI: 2004-226830/21.
P-PSDBJ:	ADK71829.
PT	New human kinases and phosphatases, useful for diagnosing, treating or
PT	preventing atherosclerosis, hypertension, AIDS, allergy, multiple
PT	sclerosis, osteoarthritis, Alzheimer's disease, Crohn's disease, cancer
PT	or hepatitis.
XX	
PS	Claim 5; SEQ ID NO 65; 347bp; English.
XX	
CC	The invention relates to a novel isolated polypeptide which is a human
CC	kinase and phosphatase (KP). The polypeptide of the invention
CC	demonstrates cardiovascular, antidiabetic, hypotensive,
CC	vasotropic, anti-inflammatory, antiangiogenic, anti-HIV, antiallergic,
CC	antibacterial, virocidic, protozoocidal and fungicide activities. The
CC	antiasthmatic, immunosuppressive, antithyroid, dermatological,
CC	osteopathic, nephrotropic, antigout, gastrointestinal, neuroprotective,
CC	antiparkinsonian, nootropic, anticonvulsant, hepatotropic, antipsoriatic,
CC	haemostatic, cytotoxic, antilipemic, antiparasitic, antimutagenic,
CC	antibacterial, virocidic, protozoocidal and fungicide activities. The
CC	kinase and phosphatase (KP) polynucleotide, polypeptides, agonists and
CC	antagonists may be useful for diagnosing, treating or preventing
CC	disorders such as cardiovascular diseases, immune system disorders,
CC	neurological disorders, disorders affecting growth and development, cell
CC	proliferative disorders and viral, bacterial, fungal, parasitic,
CC	protozoan or helminthic infections. Furthermore, the molecules of the
CC	invention may be useful for creating transgenic animals to model human
CC	disease and during gene therapy. The current sequence is that of a human
CC	KP cDNA of the invention.
XX	
SO	Sequence 2599 BP; 541 A; 786 C; 828 G; 444 T; 0 U; 0 Other;
	Query Match 97.8%; Score 2206.2; DB 12; Length 2599;
	Best Local Similarity 98.3%; Pred. No. 0;
	Matches 2253; Conservative 0; Mismatches 3; Indels 35; Gaps 1
OY	1 ATGGGCTTCTTTGTAGCTTGACCTTGACGCCCGCAGGGGCACGGGGTCCTSCAGCAATGCG 60
DB	95 ATGGGCTTCTTCCGAGCTGTGACGCCCGCAGGGGCACGGGGTCCTSCAGCAATGCG 154
OY	61 GAGGCCGACGCTTCTCTACTGAGGGCATGAGAAAGTAGATGGCCCCAGGGTCAAGACT 120
DB	155 GAGGCCGAGCTTCTGCTACTGAGGGCATGAGAAAGTAGATGGCCCCAGGGTCAAGACT 214
OY	121 GACAGGGAGTAGACAGGACGTGCTTCAACAATGTCCCTGACAGGAAGTAGGGGCCAGAGC 180
DB	215 GACAGGGAGTAGACAGGACGTGCTTCAACAATGTCCCTGACAGGAAGTAGGGGCCAGAGC 274
OY	181 CGGGCCATCACGCCCTTGACAGGCCCATCATGATCTTGSGGTGAGATCAACGCAAACT 240
DB	275 CGGGCCATCACGCCCTTGACAGGCCCATCATGATCTTGSGGTGAGATCAACGCAAACT 334
OY	241 GAGGGCTTGAGCCGCTTGCTGGGCGACAGCAGCAGAGATCTGAATCTCAGGCGCCCTGAGC 300
DB	335 GAGGGCTTGAGCCGCTTGCTGGGCGACAGCAGCAGAGATCTGAATCTCAGGCGCCCTGAGC 394
OY	301 AAGCTGAGCCGTCTCATCCGGGAAACGGAGAGAGCTTGCAAGAGCTTACAGGAGAGAGG 360
DB	395 AAGCTGAGCCGTCTCATCCGGGAAACGGAGAGAGCTTGCAAGAGCTTACAGGAGAGAGG 454
OY	361 CAGCAGCTGACAGAGAGACTCAACCAAGACCACAGCCAGGACATTGAGAACTGTAAGAGC 420
DB	455 CAGCAGCTGACAGAGAGACTCAACCAAGACCACAGCCAGGACATTGAGAACTGTAAGAGC 514
OY	421 CAGTACCGAGCTTGACAGGGGACAGTGGCCAGAGCCGCAAGTACCAAGAGGCGAGC 480

Db	515	CAGTACCGAGCTTGTGGCACGGGACAGTCCCAAGCCAAAGTACCAAGAGCCACGC	574
Qy	481	AAAGCAAGAACCGTGAACAAGGCCAAGACAAGTATGTGCCAGCTGTGAAGCTCTT	540
Db	575	AAAGCAAGAACCGTGAACAAGGCCAAGACAAGTATGTGCCAGCTGTGAAGCTCTT	634
Qy	541	GCTCACCAACAACCGTATATGTGTGGGGCTGTGGGCTGTGGCAGCTACACCAACGACAC	600
Db	635	GCTCACCAACAACCGTATATGTGTGGGGCTGTGGGCTGTGGCAGCTACACCAACGACAC	694
Qy	601	CACCAAGCTCTGTGGCCCGGCTGTGGGTCATCTGACAGACCTTGACAAGAAATGGCT	660
Db	695	CACCAAGCTCTGTGGCCCGGCTGTGGGTCATCTGACAGACCTTGACAAGAAATGGCT	754
Qy	661	TGCATCTCTGAAGAGATCTCTGACAGAAATACCTTGAGATTTAACAGCTGTGACAGATAG	720
Db	755	TGCATCTCTGAAGAGATCTCTGACAGAAATACCTTGAGATTTAACAGCTGTGACAGATAG	814
Qy	721	GTGTGTGCATTCACCGGGAATGGCTGACAGCTGTGCTGCCGATCAGCTTGAAGGTAG	780
Db	815	GTGTGTGCATTCACCGGGAATGGCTGACAGCTGTGCTGCCGATCAGCTTGAAGGTAG	874
Qy	781	TACCAAGGCTTCTGTCCGACAGTATGGGTTCGGCACTTGAACCTTCCACCTGTGTCACTTC	840
Db	875	TACCAAGGCTTCTGTCCGACAGTATGGGTTCGGCACTTGAACCTTCCACCTGTGTCACTTC	934
Qy	841	GATGAGTCACTGCTTGAAGAGGTGAACCGCTGAGACCTGGGGAGCTCAGCTGAACGAG	900
Db	935	GATGAGTCACTGCTTGAAGAGGTGAACCGCTGAGACCTGGGGAGCTCAGCTGAACGAG	994
Qy	901	CTGACTGTGAGAGCGGTGACACACGCTGACCTTCAAGTACAGATGAGCTGGCTGTGGCC	960
Db	995	CTGACTGTGAGAGCGGTGCTGCAACGCTGACCTTCAAGTACAGATGAGCTGGCTGTGGCC	1054
Qy	961	ACCGAGATGTGTTCAGGCGGCAAGAAATGTGTACGACGCTGCAACAGAGCTCCGAAAT	1020
Db	1055	ACCGAGATGTGTTCAGGCGGCAAGAAATGTGTACGACGCTGCAACAGAGACTCCGAAAT	1114
Qy	1021	GAAGAGAGAAACACCCACCCCGGAGGGGGTGCAGCTGTGGCAAGGCAAGTGTG	1080
Db	1115	GAAGAGAGAAACACCCACCCCGGAGGGGGTGCAGCTGTGGCAAGGCAAGTGTG	1174
Qy	1081	CAAGAGCACTGCAAGGGGCTCAGGTTAGCGCTGTGACAGCAGGCCAAGCTCAGGCCAG	1140
Db	1175	CAAGAGCACTGCAAGGGGCTCAGGTTAGCGCTGTGACAGCAGGCCAAGCTCAGGCCAG	1234
Qy	1141	CAGAGATTGTCTGCAAGACCAAGCTTGAGACACTTGGGCCCCGGAGCCCCGCTGTGCTG	1200
Db	1235	CAGAGATTGTCTGCAAGACCAAGCTTGAGACACTTGGGCCCCGGAGCCCCGCTGTGCTG	1294
Qy	1201	CTCCGACAGATGACCGGCACTCCACGTCGTCCTG-----	1235
Db	1295	CTCCGACAGATGACCGGCACTCCACGTCGTCCTG-----	1354
Qy	1236	-----GAGCAGAGAGCAGAGGGGGGAAGACACCAACGCTGAGATCTCTTAAGA	1285
Db	1355	CTTGGGCTCAGAGAGCAGAGGGGGGAAGACACCAACGCTGAGATCTCTTAAGA	1414
Qy	1286	GCCACATCTCAGAAATCTTCCGCCCAAGTTCTGSACTCTGTACCCTGAAAGGGAGG	1345
Db	1415	GCCACATCTCAGAAATCTTCCGCCCAAGTTCTGSACTCTGTACCCTGAAAGGGAGG	1474
Qy	1346	GCTTTCCTAGAAATCTTGTGCATTCGACCACTAATGAGACCCAGAGGCCCTCACA	1405
Db	1475	GCTTTCCTAGAAATCTTGTGCATTCGACCACTAATGAGACCCAGAGGCCCTCACA	1534
Qy	1406	AGAAAGTGTGTGTCTCTGCAACAGGGCTGTGCCCAAGACAAGTGTGGTGTGAAACATG	1465
Db	1535	AGAAAGTGTGTGTCTCTGCAACAGGGCTGTGCCCAAGACAAGTGTGGTGTGAAACATG	1594
Qy	1466	AGGACTGTGTGTGGGTGACAGATTTGACCGGGGAATTTTGGCGAATGTTCAAGCGAC	1525

Db	1535	AGGACCTG6TGTGGGTGAGCAGATTTGGA	CGGGGAACTTTGGCGGAAGTCTTCA	CGGCAC	1634	
Qy	1526	GCCTGCGAGCCGACAACA	CCCTTGATGGCGGTGAAGTCTTGT	CGAGAG	CAGCTCCACCTG	1585
Db	1655	GCCTGCGAGCCGACAACA	CCCTGGTGGCGGTGAAGTCTTGT	CGAGAG	CAGCTCCACCTG	1714
Qy	1586	ACCTCAGGGCCAAAGTTTCTTAC	AGAGAGCGAGGATCTCGAAGCAGTAC	AGCCACCCCAACA		1645
Db	1715	ACCTCAAGGCCAAAGTTTCTTAC	AGAGAGCGAGGATCTCGAAGCAGTAC	AGCCACCCCAACA		1774
Qy	1646	TGGTGGCTCATTTGGTGTCTGCA	CCCAAGAGAGCCCATATACATGTCAT	TGAGAGCTTG		1705
Db	1775	TGGTGGCTCATTTGGTGTCTGCA	CCCAAGAGAGCCCATATACATGTCAT	TGAGAGCTTG		1833
Qy	1706	TGCAGGGGGGCGAATTCTCTGA	CTTCTCTCCGACGAGGGGGCCCGCT	TGCGGGTGAAGA		1765
Db	1835	TGCAGGGGGGCGAATTCTCTGA	CTTCTCTCCGACGAGGGGGCCCGCT	TGCGGGTGAAGA		1894
Qy	1766	CTCTGCTGCAGATGGTGGGGGATG	CAGTCTGTCGATGAGATACCTGGAGACA	GTGCT		1823
Db	1895	CTCTGCTGCAGATGGTGGGGGATG	CAGTCTGTCGATGAGATACCTGGAGACA	GTGCT		1951
Qy	1826	GCATTCACCCGGGACCTGCTGCTG	GGAACCTGCGTGTGACAGAGAAGAA	TTGCTCTGAAGA		1885
Db	1955	GCATTCACCCGGGACCTGCTGCTG	GGAACCTGCGTGTGACAGAGAAGAA	TTGCTCTGAAGA		2014
Qy	1886	TCAGTGACTTTGGGATGTCCCGA	AGAGAGCCGATGGGGTCTATG	CAGCCTCAGGGGCTC		1945
Db	2015	TCAGTGACTTTGGGATGTCCCGA	AGAGAGCCGATGGGGTCTATG	CAGCCTCAGGGGCTC		2074
Qy	1946	TCAGACAAGTCCCGGTGAATGGA	ACCGCACTGAGGCCCTTAACTAC	CGCCGTACTTCT		2005
Db	2075	TCAGACAAGTCCCGGTGAATGGA	ACCGCACTGAGGCCCTTAACTAC	CGCCGTACTTCT		2134
Qy	2006	CCGAAAGGAGTGTGAGAGCTTTGG	CACTTGGCTCTGCGAAGACCTTCA	GCCCTGGGGGCTC		2065
Db	2135	CCGAAAGGAGTGTGAGAGCTTTGG	CACTTGGCTCTGCGAAGACCTTCA	GCCCTGGGGGCTC		2194
Qy	2066	CCCCCTATCCCAACTCAGCAAT	CAGACAGACACCGAGATTGTG	AGAGAGGGGGGCGCTC		2125
Db	2195	CCCCCTATCCCAACTCAGCAAT	CAGACAGACACCGAGATTGTG	AGAGAGGGGGGCGCTC		2254
Qy	2126	TGCCCTGCCAGAGCTGTCTTGAT	GCCGTGTTCAAGCTTCATGAGACAG	TGCTGGGCTT		2185
Db	2255	TGCCCTGCCAGAGCTGTCTTGAT	GCCGTGTTCAAGCTTCATGAGACAG	TGCTGGGCTT		2314
Qy	2186	ATTGAGCCTGGGCAACGGCCGCA	GCTTTCAGACACATCTACAGAG	ACTCGACAGCATCCGAA		2245
Db	2315	ATTGAGCCTGGGCAACGGCCGCA	GCTTTCAGACACATCTACAGAG	ACTCGACAGCATCCGAA		2374
Qy	2246	AGCGGCAATCGG	2256			
Db	2375	AGCGGCAATCGG	2385			
RESULT 5						
AAS95001						
ID	AAS95001 standard; DNA; 2889 BP.					
XX	AAS95001;					
XX						
DT	14-FEB-2002 (first entry)					
XX						
DE	Human DNA sequence #256 expressed during foam cell differentiation.					
XX						
XX	Human; foam cell differentiation; atherosclerosis; cerebral stroke;					
KW	cardiovascular disorder; coronary artery disease; gene therapy; ds.					
XX						
OS	Homo sapiens.					
XX						
PN	WO200177389-A2.					
XX						
PD	18-OCT-2001.					

XX 04-APR-2001; 2001WO-US011128.
XX
XX
XX 05-APR-2000; 2000US-0195106P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Shiffman D, Somogyi R, Lawn R, Seilhamer JU, Porter GU, Mikita T,
XX Tel J;
XX WPI, 2002-010925/01.
XX
XX Composition useful for diagnosis of conditions, disorders or diseases
XX associated with atherosclerosis, comprises several polynucleotides that
XX are differentially expressed in foam cell development.
XX
XX Claim 1; Page 298-299; 315pp; English.
XX
XX The present invention relates to the isolation of human polynucleotide
XX sequences that are differentially expressed during foam cell
XX differentiation. The polynucleotide sequences of the invention or a
XX composition comprising these polynucleotides are useful as a high
XX throughput method for detecting altered expression of one or more
XX polynucleotides in a sample. The polynucleotides can be used in the
XX diagnosis of disorders associated with foam cell development such as
XX atherosclerosis, cerebral stroke, and cardiovascular disorders such as
XX coronary artery disease. The polynucleotide sequences can also be used as
XX PCR primers and probes. The polynucleotides of the invention are also
XX useful in gene therapy. AAS94746-AAS95021 represent the human
XX polynucleotide sequences of the invention which are differentially
XX expressed during foam cell differentiation
XX
SQ Sequence 2889 BP; 600 A; 873 C; 929 G; 487 T; 0 U; 0 Other;
Query Match 89.7%; Score 2023.4; DB 6; Length 2889;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 2255; Conservative 0; Mismatches 1; Indels 211; Gaps 2;
QY 1 ATGGGGCTTCTCTCTGAGCTGTGAGCGCCAGAGGCGACGGGGTCTCTGACGAATGACAG 60
DB 188 ATGGGCTTCTCTCTGAGCTGTGAGCGCCAGAGGCGACGGGGTCTCTGACGAATGACAG 247
QY 61 GAGGCCGAGCTTCTGTCTACTGAGGGCATGAGAAATGATGAGTGGCCAGCGGGTCAAAGAT 120
DB 248 GAGGCCGAGCTTCTGTCTACTGAGGGCATGAGAAATGATGAGTGGCCAGCGGGTCAAAGAT 307
QY 121 GACAGGAGATGATGACAGACTGCTTACACACATGCTCCCTGACAGACAGTGGGGCCAGAGC 180
DB 308 GACAGGAGATGATGACAGACTGCTTACACACATGCTCCCTGACAGACAGTGGGGCCAGAGC 367
QY 181 CGGGCCATGACGCTCTGACAGCGCCCATCAATCAATCTGAGCTGAGATCAACAGCAAACT 240
DB 368 CGGGCCATGACGCTCTGACAGCGCCCATCAATCAATCTGAGCTGAGATCAACAGCAAACT 427
QY 241 GA-GGGCTTGAAGCGCTTGTCTGCGGAGACGCGAGAGATCTGAACCTCAGGGCCCTTGA 299
DB 428 GAGGGCTTGAAGCGCTTGTCTGCGGAGACGCGAGAGATCTGAACCTCAGGGCCCTTGA 487
QY 300 CAAAGCTGAGCTTCTCTATCCGGGAGAGCGGACGCTTTCGCAAGACCTTACAGGAGAGTG 359
DB 488 CAAAGCTGAGCTTCTCTATCCGGGAGAGCGGACGCTTTCGCAAGACCTTACAGGAGAGTG 547
QY 360 GCAGAGCTGACAGAGAGCTCAACAAGACCAAGCAGAGCAGTTCAGAGAGTGAAGAG 419
DB 548 GCAGAGCTGACAGAGAGCTCAACAAGACCAAGCAGAGCAGTTCAGAGAGTGAAGAG 607
QY 420 CCAAGTACCAAGCTCTGACAGGAGACAGTGCACCAAGCCAAAGCAGTTCAGAGAGCAG 479
DB 608 CCAAGTACCAAGCTCTGACAGGAGACAGTGCACCAAGCCAAAGCAGTTCAGAGAGCAG 667
QY 480 CAAAGACCAAGAGACCTGACAGAGCCAAAGACCAAGTATGTGCGAGCTGTGAGAGCTTT 539
DB 668 CAAAGACCAAGAGACCTGACAGAGCCAAAGACCAAGTATGTGCGAGCTGTGAGAGCTTT 727

QY 540 TECTACCAACACCGCTATGTGCTGGGCTGGCGGGCTGGCAGCTTACACACAGACCA 599
DB 728 TECTACCAACACCGCTATGTGCTGGGCTGGCGGGCTGGCAGCTTACACACAGACCA 787
QY 600 CCAACGAGCTCTGCTGCGCCGCGCTGTGCTGCTGACAGACCTGACAGAGATGGC 659
DB 788 CCAACGAGCTCTGCTGCGCCGCGCTGTGCTGCTGACAGACCTGACAGAGATGGC 847
QY 660 TTGCATCTGAGAGAGATCTGACAGAAATCACTGGAGATTAGACAGCTGTGAGAGATGA 719
DB 848 TTGCATCTGAGAGAGATCTGACAGAAATCACTGGAGATTAGACAGCTGTGAGAGATGA 907
QY 720 GGTGTGGCTCATTCACCGGAGATGCTGACAGCTGCTGCGCATCCAGCTGAGAGCTGA 779
DB 908 GGTGTGGCTCATTCACCGGAGATGCTGACAGCTGCTGCGCATCCAGCTGAGAGCTGA 967
QY 780 GTACCAAGGCTTCTGCGACAGTATGGGTCCGACCTGACAGCTCCACCTGTGTACGTT 839
DB 968 GTACCAAGGCTTCTGCGACAGTATGGGTCCGACCTGACAGCTCCACCTGTGTACGTT 1027
QY 840 CGATGAGTCACTGTTGAGAGAGGTGAACCGGTGAGCGCTGGGAGCTCCAGCTGAACGA 899
DB 1028 CGATGAGTCACTGTTGAGAGAGGTGAACCGGTGAGCGCTGGGAGCTCCAGCTGAACGA 1087
QY 900 GCTGACTGTGAGAGCGTGCACACACGCTGACCTCAGTACAGATGAGCTGTGCGC 959
DB 1088 GCTGACTGTGAGAGCGTGCACACACGCTGACCTCAGTACAGATGAGCTGTGCGC 1147
QY 960 CACCGAGATGTGTTTCAAGCGGCGAGAGATGTTAGCGACGCTGCAACAGAGTCCGGAA 1019
DB 1148 CACCGAGATGTGTTTCAAGCGGCGAGAGATGTTAGCGACGCTGCAACAGAGTCCGGAA 1207
QY 1020 TGAAGAGAGAAACCCACCCCGGAGGCGGTGACGTGCTGGGAGAGGCAATGCT 1079
DB 1208 TGAAGAGAGAAACCCACCCCGGAGGCGGTGACGTGCTGGGAGAGGCAATGCT 1267
QY 1080 GCAAGAGACCTGCAAGGGGCTGCAAGGTACGCTGTGACAGCCAGCCAACTGAGGCCA 1139
DB 1268 GCAAGAGACCTGCAAGGGGCTGCAAGGTACGCTGTGACAGCCAGCCAACTGAGGCCA 1327
QY 1140 GCAAGAGTGTGCTGACAGACCAAGTGTGACACTGTGGCCCGCGAGACCCCGCTGTGCT 1199
DB 1328 GCAAGAGTGTGCTGACAGACCAAGTGTGACACTGTGGCCCGCGAGACCCCGCTGTGCT 1387
QY 1200 GCTCCTGAGAGATGACCGGCACTCCAGTGTCTGTGAGAGAGAGAGGAGGAG 1259
DB 1388 GCTCCTGAGAGATGACCGGCACTCCAGTGTCTGTGAGAGAGAGAGGAGGAG 1447
QY 1260 GACACCCAGCTGTGAGATCTTAAAGACCACTTGAAGAAATCTTCGCGCCAAATTTCTC 1319
DB 1448 GACACCCAGCTGTGAGATCTTAAAGACCACTTGAAGAAATCTTCGCGCCAAATTTCTC 1507
QY 1320 G----- 1320
DB 1508 GCTCCTTCCACCGCTGACAGCTAATCCGAGAGGTGAGAAAGCCCTGATGAGCAGCTGTG 1567
QY 1321 ----- 1320
DB 1568 GTACCAAGGAGCCATCCGAGGGCAGAGGTGGCTGAGCTGTGTGACACTTGTGGGACTT 1627
QY 1321 ----- 1320
DB 1628 CCGTGTGGGAGAGAGCCAGGGCAAGAGATACGTGTGTGTGTGTGTGTGTGTGTGT 1687
QY 1321 -----AACTGTACCACTGAGAGAGAGAGCTT 1349
DB 1688 GCCCGGCACTTCATCATCATCATCTTGTGATTAACCTGTACCGACTGTGAGAGAGAGCTT 1747
QY 1350 TCTTACATTCCTTGTGCTATGACACCACTTACAGACCCAGCAGCCCTTACCAAGAA 1409
DB 1748 TCTTACATTCCTTGTGCTATGACACCACTTACAGACCCAGCAGCCCTTACCAAGAA 1807

QY 1410 GAGTGGTGTTCCTGCAACAGGGCTGTGCCCAAGACAAAGTGGGTGTGAACCATAGGA 1469
 DB 1808 GAGTGGTGTTCCTGCAACAGGGCTGTGCCCAAGACAAAGTGGGTGTGAACCATAGGA 1867
 QY 1470 CCTGTGTGTGGTGAACAGATTTGAGCGGGGAACTTTGGCGAAGTGTTCAGCGGAGCCCT 1529
 DB 1868 CCTGTGTGTGGTGAACAGATTTGAGCGGGGAACTTTGGCGAAGTGTTCAGCGGAGCCCT 1927
 QY 1530 GCGAGCGCAACACACCTTGTGGTGAAGTCTTGTGAGAGACGCTCCACCTGACCT 1589
 DB 1928 GCGAGCGCAACACACCTTGTGGTGAAGTCTTGTGAGAGACGCTCCACCTGACCT 1987
 QY 1590 CAAGGCGCAAGTTCTACAGGAAGCGAGATTCCTGAAGCGATACAGCCCAACCTCG 1649
 DB 1988 CAAGGCGCAAGTTCTACAGGAAGCGAGATTCCTGAAGCGATACAGCCCAACCTCG 2047
 QY 1650 GCGTCTCATTTGTGTGTGCAACCGAAGAGCCCATCTACATGATGAGAGCTTGTGCA 1709
 DB 2048 GCGTCTCATTTGTGTGTGCAACCGAAGAGCCCATCTACATGATGAGAGCTTGTGCA 2107
 QY 1710 GGGGGGCGACTTCTCTGACCTTCTCCGACGAGAGGGGGCCGCTCGGGGTGAAGACTCT 1769
 DB 2108 GGGGGGCGACTTCTCTGACCTTCTCCGACGAGAGGGGGCCGCTCGGGGTGAAGACTCT 2167
 QY 1770 GCTGCGAGATGCTGGGGAGTGCAGCTGTGCAATGAGTACCTGAGAGAGAGTGTGCAT 1829
 DB 2168 GCTGCGAGATGCTGGGGAGTGCAGCTGTGCAATGAGTACCTGAGAGAGAGTGTGCAT 2227
 QY 1830 CCACCGGAGACTTGGCTGTGCAAGTCTGTGAGAGAGAGATGTCTTGAAGATCAG 1889
 DB 2228 CCACCGGAGACTTGGCTGTGCAAGTCTGTGAGAGAGAGATGTCTTGAAGATCAG 2287
 QY 1890 TGACTTTGGAGATGCCCGAGAGAGAGCCGATGATGAGAGAGAGTCTGAGAGAGTCTGAG 1949
 DB 2288 TGACTTTGGAGATGCCCGAGAGAGAGCCGATGATGAGAGAGAGTCTGAGAGAGTCTGAG 2347
 QY 1950 ACAAGTCCCCCGTGAAGTGAACCCGACCTGAGGCCCTTAATAGGCCGCTACTCTCCGA 2009
 DB 2348 ACAAGTCCCCCGTGAAGTGAACCCGACCTGAGGCCCTTAATAGGCCGCTACTCTCCGA 2407
 QY 2010 AAGCGAGTGTGAGAGCTTTGGCATCTTGTGCTGGAGAGACCTTGAAGCTGGGGGCTCCCG 2069
 DB 2408 AAGCGAGTGTGAGAGCTTTGGCATCTTGTGCTGGAGAGACCTTGAAGCTGGGGGCTCCCG 2467
 QY 2070 CTATCCCACTCAGCAATCAGACACGCGAGTTTGTGAGAGAGGGGGGCGCTGTGCC 2129
 DB 2468 CTATCCCACTCAGCAATCAGACACGCGAGTTTGTGAGAGAGGGGGGCGCTGTGCC 2527
 QY 2130 CTGCCAGAGCTGTGCTGATGATGCCGTTCAGGCTCATGAGACAGTGTGGGCTCATGGA 2189
 DB 2528 CTGCCAGAGCTGTGCTGATGATGCCGTTCAGGCTCATGAGACAGTGTGGGCTCATGGA 2587
 QY 2190 GCGTGGGAGAGGCGGCGAGCTTTCAGACACATCTACAGAGAGCTGAGAGCATCCGAAAGCG 2249
 DB 2588 GCGTGGGAGAGGCGGCGAGCTTTCAGACACATCTACAGAGAGCTGAGAGCATCCGAAAGCG 2647
 QY 2250 GCATCGG 2256
 DB 2648 GCATCGG 2654
 RESULT 6
 ADR39816 standard; cDNA; 2623 BP.
 ADR39816;
 18-NOV-2004 (first entry)
 Human kinase and phosphatase KPP-43 encoding cDNA SEQ ID NO:89.
 human: kinase and phosphatase protein; KPP: enzyme; cytosolic;
 antiarteriosclerotic; anticonvulsant; nootropic; neuroprotective;

KW cerebroprotective; anti-HIV; antiallergic; anti-inflammatory;
 KW thymimetic; gene therapy; cell proliferative disorder; cancer;
 KW atherosclerosis; neurological disorder; epilepsy; Huntington's disease;
 KW stroke; immune disorder; inflammatory disorder; AIDS; allergy;
 KW developmental disorder; Hypothyroidism; Cushing's syndrome; infection;
 KW KPP-43; single nucleotide polymorphism; SNP; gene; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 95..2389
 FT /*tag= a
 FT /*product= "kinase and phosphatase KPP-43"
 FT variation replace(109,C)
 FT /*tag= b
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 PD MO2004074453-AZ.
 PD 02-SEP-2004.
 PF 20-FEB-2004; 2004MO-US005092.
 PR 20-FEB-2003; 2003US-0449059P.
 PR 19-MAR-2003; 2003US-0456932P.
 PR 28-MAR-2003; 2003US-0458844P.
 PR 09-APR-2003; 2003US-0461678P.
 PR 17-APR-2003; 2003US-0463937P.
 PA (INCY-) INCYTE CORP.
 PI Ramkumar J, Margulis JP, Swarnakar A, Chawla NK, Tran UK;
 PI Becha SD, Lee SY, Hafalia AJA, Richardson TW, Khare R, Jiang X;
 PI Jackson AA, Yang J, Gorvad AE;
 DR WPI: 2004-635568/61.
 DR P-PADB; ADR39770.
 XX
 XX
 PT New human kinases and phosphatases (KPP) for diagnosing, treating and
 PT preventing diseases or conditions associated with aberrant KPP expression
 PT e.g. cancer, acquired immunodeficiency syndrome, epilepsy, or infections.
 XX
 PS Claim 5; SEQ ID NO 89; 299pp; English.
 XX
 XX The present sequence encodes the human kinase and phosphatase protein
 CC (KPP), designated KPP-43. The human KPP sequences from the present
 CC invention have cytosolic, antiarteriosclerotic, anticonvulsant,
 CC nootropic, neuroprotective, cerebroprotective, anti-HIV, antiallergic,
 CC antiinflammatory and thymimetic activities, and can be used in gene
 CC therapy. The human KPP proteins and polynucleotides can be used in
 CC diagnosing, treating and preventing diseases or conditions associated
 CC with the decreased expression or overexpression of KPP, such as cell
 CC proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
 CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
 CC allergies) and developmental (e.g. Hypothyroidism, Cushing's syndrome),
 CC disorders, or infections. They can also be used in assessing the effects
 CC of exogenous compounds on the expression of nucleic acid and amino acid
 CC sequences of KPP. The KPP or its fragments are useful in screening
 CC compounds for effectiveness as agonist or antagonist of the polypeptides,
 CC or in altering the expression of the target polynucleotide and compounds
 CC that specifically bind to or modulate the activity of the polypeptide.
 XX
 SQ Sequence 2623 BP; 541 A; 786 C; 840 G; 456 T; 0 U; 0 Other;
 XX
 Query Match 74.1%; Score 1671.6; DB 13; Length 2623;
 Best Local Similarity 84.3%; Pred. No. 1.1e-305;
 Matches 2078; Conservative 0; Mismatches 4; Indels 384; Gaps 2;
 QY 1 ATGGGCTTCTCTTCTGAGCTGTGACGCCCCCAGCGGCGCTCTGACGCAATGACG 60
 DB 95 ATGGGCTTCTCTTCTGAGCTGTGACGCCCCCAGCGGCGCTCTGACGCAATGACG 154
 QY 61 GAGGCGGAGCTTGTCTACTGAGGCGCATGAGAAATGAGATGGCCAGCGGCTCAAGAGT 120

155 GAGCGGAGCTTGTCTACTGAGGCGATGAGAAATGATGCGCCAGCGGGCTCAAGAGT 214
121 GACAGGAGATATATACAGAGCTGCTTCAACACATGTCCCTGACAGACAGTGGGGCCAGAGC 180
215 GACAGGAGATATGACAGAGCTGCTTCAACACATGTCCCTGACAGACAGTGGGGCCAGAGC 274
181 CGGGCCATCAGCGCCCTGACAGCGCCCATCAGTCACTGCGGCTGATGATCACAGCCAAACT 240
275 CGGGCCATCAGCGCCCTGACAGCGCCCATCAGTCACTGCGGCTGATGATCACAGCCAAACT 305
241 GAGGCGCTGAGCGGCTTGTCTGCGGACAGACAGAGAGATCTGAATCAAGGCGCCCTGAGC 300
306 ----- 305
301 AAGCTAGCTGCTCATCCGGGAAACGGACAGAGCTTCGAAAGACTTAAGCCAGACAGTGG 360
306 ----- 305
361 CAGCAGCTGACAGAGAGCTCACCAAGACCCACAGCCAGGACATTTGAGAGTGAAGAGC 420
306 -----AGACCCACAGCCAGGACATTTGAGAGTGAAGAGC 340
421 CAGTACCGAGCTCTGACAGGAGCAGTGCCTCAAGCCAAAGCAAGTACAGAGGCGCAGC 480
341 CAGTACCGAGCTCTGACAGGAGCAGTGCCTCAAGCCAAAGCAAGTACAGAGGCGCAGC 400
481 AAGACAAAGAACCGGTGACAAAGGCCAAAGACAAATATGTGCGAGCCCTGTGGAAGCTCTTT 540
401 AAGACAAAGAACCGGTGACAAAGGCCAAAGACAAATATGTGCGAGCCCTGTGGAAGCTCTTT 460
541 GCTCACCAACACCGCTATGTGCTGCGGCGTGCAGGCTGCGAGCTACACCAAGCAGCAC 600
461 GCTCACCAACACCGCTATGTGCTGCGGCGTGCAGGCTGCGAGCTACACCAAGCAGCAC 520
601 CACCACTCTGCTGCTGCGGCGTGCAGGCTGCGAGCTGCACTGACAGGACCTGCAAGAGAGTGCCT 660
521 CACCACTCTGCTGCTGCGGCGTGCAGGCTGCGAGCTGCACTGACAGGACCTGCAAGAGAGTGCCT 580
661 TGCATCTGAGAGAGATCTCTGAGAGAAATACCTGGAATTTAGCGAGCCCTGTGAGAGATGAG 720
581 TGCATCTGAGAGAGATCTCTGAGAGAAATACCTGGAATTTAGCGAGCCCTGTGAGAGATGAG 640
721 GTGTGGCCATTACCGGGAGATGAGCTGACAGTGCCTGCGCACTCCAGCTTGAAGCTGAG 780
641 GTGTGGCCATTACCGGGAGATGAGCTGACAGTGCCTGCGCACTCCAGCTTGAAGCTGAG 700
781 TACCAAGGCTTCTGCGACAGTATGGGTCCGACCTGACGCTCCACCTGTGTCAAGTTTC 840
701 TACCAAGGCTTCTGCGACAGTATGGGTCCGACCTGACGCTCCACCTGTGTCAAGTTTC 760
841 GATGAGCTCATGCTTGAAGAGGGTGAACCGCTGAGGCTGCGGAGAGCTCCAGCTGAAGAG 900
761 GATGAGCTCATGCTTGAAGAGGGTGAACCGCTGAGGCTGCGGAGAGCTCCAGCTGAAGAG 820
901 CTGACTGTGAGAGCGTGCAGCAACCGCTGACCTCAGTGAACAGATGAGTGCCTGTGAGC 960
821 CTGACTGTGAGAGCGTGCAGCAACCGCTGACCTCAGTGAACAGATGAGTGCCTGTGAGC 880
961 ACCGAGATGTGTTCAGGCGGACAGAGATGTTTACGAGCTGCAACAGAGCTCCGGAAT 1020
881 ACCGAGATGTGTTCAGGCGGACAGAGATGTTTACGAGCTGCAACAGAGCTCCGGAAT 940
1021 GAAGAGAGAAACACCCACCCCGGAGAGGGGTGACAGTGTGCGCAAGAGAGAGAGAGCTG 1080
941 GAAGAGAGAAACACCCACCCCGGAGAGGGGTGACAGTGTGCGCAAGAGAGAGAGAGCTG 1000
1081 CAAGAAAGCACTGACAGGGGCTGACAGGTGAGCTGTGACAGGCAAGAGCTGACAGGCGCAG 1140
1001 CAAGAAAGCACTGACAGGGGCTGACAGGTGAGCTGTGACAGGCAAGAGCTGACAGGCGCAG 1060
1141 CAGAGTGTGTCAGACCAAGCTGAGACACTTGGGCCCGGCGAGGCCCGCTGTGCTG 1200

1061 CAGAGTGTCTCAGACCAAGCTGAGACCTGCGGCCCGGCGAGGCCCGCTGTGCTG 1120
1201 CTCTGACAGATGACCGCACTTCCAGTGTCTCTCGAGACAGAGAGAGAGGGGAAAG 1260
1121 CTCTGACAGATGACCGCACTTCCAGTGTCTCTCGAGACAGAGAGAGAGGGGAAAG 1180
1261 ACACCCACGCTGAGAGATCTTAAAGCCACATCTCAGAAATCTTCCGCCCAAGTTCTG 1320
1181 ACACCCACGCTGAGATCTTAAAGCCACATCTCAGAAATCTTCCGCCCAAGTTCTG 1240
1321 ----- 1320
1241 CTCTTCCACCGCTGACCTCATTCGAGGTGCAGAGGCCCTGTGATAGACAGCTTGG 1300
1321 ----- 1320
1301 TACACAGGGGCCATCCGAGGGCAGAGGTGGCTGAGCTGTGTGACACTTGGGGACTTC 1360
1321 ----- 1320
1361 CTGTGCGGAGAGGCCAGGGCAAGCAGAGTACGTGTGCGGTGCTGTGGATGTCTG 1420
1321 -----AACCTGTACCGACTGGAAGGGAGGCTTT 1350
1421 CCGCGCACTTATCATTCAGATCTTGAATACCTGTACCGACTGGAAGGGAGGCTTT 1480
1351 CCTAGCATTCCTTGTGCTATGACCACTTACTGAGACCCAGCAGCCCTTCAACCAAGAG 1410
1481 CCTAGCATTCCTTGTGCTATGACCACTTACTGAGACCCAGCAGCCCTTCAACCAAGAG 1540
1411 AGTGTGTGTCTGTGACAGGGCTGTGCCAAGACAAAGTGGTGTGAACCATGAGAC 1470
1541 AGTGTGTGTGTCTGTGACAGGGCTGTGCCAAGACAAAGTGGTGTGAACCATGAGAC 1600
1471 CTGTGTGTGTGTGAGAGATTTGAGCGGGGAACTTTGGGAAAGTGTTCAGCGGACGCTG 1530
1601 CTGTGTGTGTGTGAGAGATTTGAGCGGGGAACTTTGGGAAAGTGTTCAGCGGACGCTG 1660
1531 CGAGCCGACAAACCTGTGTGCGGAGTCTTGTGAGAGACGCTCCACCTGACCTC 1590
1661 CGAGCCGACAAACCTGTGTGCGGAGTCTTGTGAGAGACGCTCCACCTGACCTC 1720
1591 AAGGCCAAGTTTCTACAGAAAGGAGAGATCTGAAACAGTACAGCAACCCCAACATGCTG 1650
1721 AAGGCCAAGTTTCTACAGAAAGGAGAGATCTGAAACAGTACAGCAACCCCAACATGCTG 1780
1651 CGTCTATGTGTGTCTGACACCCAGAAAGCAGCCCATTAATGCTCATGAGAGCTTGTGAG 1710
1781 CGTCTATGTGTGTCTGACACCCAGAAAGCAGCCCATTAATGCTCATGAGAGCTTGTGAG 1840
1711 GGGGCGCACTTCTGACCTTCTCCGACAGAGGGGCGGCTGTGCGGATGAAGACTTG 1770
1841 GGGGCGCACTTCTGACCTTCTCCGACAGAGGGGCGGCTGTGCGGATGAAGACTTG 1900
1771 CTGACAGATGTGGGAGTSCAGCTGTGACATGAGATACCTGAGAGCAAGTGTGATC 1830
1901 CTGACAGATGTGGGAGTSCAGCTGTGACATGAGATACCTGAGAGCAAGTGTGATC 1960
1831 CACCGGACCTGTGCTCTGGAACCTGCTGTGACAGAGAAAGAAATCTTGAAGATCAGT 1890
1961 CACCGGACCTGTGCTCTGGAACCTGCTGTGACAGAGAAAGAAATCTTGAAGATCAGT 2020
1891 GACTTGGAGATGCCGAGAGAAAGCGATGGGGTCTATGACAGCTCAGGGGCGCTCAGA 1950
2021 GACTTGGAGATGCCGAGAGAAAGCGATGGGGTCTATGACAGCTCAGGGGCGCTCAGA 2080
1951 CAAATCCCGTGAAGTGAACCGCACTGAGGCCCTTAACCTACGAGCGCTTACTCTCGAA 2010
2081 CAAATCCCGTGAAGTGAACCGCACTGAGGCCCTTAACCTACGAGCGCTTACTCTCGAA 2140
2011 AGCAGCTGTGAGCTTTGTGCACTTGTCTGTGGAACCTTCAAGCTGTGGGGGCTCCCC 2070
2141 AGCAGCTGTGAGCTTTGTGCACTTGTCTGTGGAACCTTCAAGCTGTGGGGGCTCCCC 2200

QY 1081 CAGAGCACTGAGGGGCTGAGAGGCTGTGAGAGCCAGGCTGAGGCCAG 1140
 DB 1140 CAGAGGCAATACAGAGGTTGAGATGACCTGTGAGAGCAAGCTTCAGGCCAG 1199
 QY 1141 CAGAGTTGTGAGAGCAAGCTGAGAGCACTGTGGGCCCCGAGGCCCTGTGTG 1200
 DB 1200 CAGGAATCTGTGAGAGCAAGATGAGAGCTCGGACAGGGGCGAGGCCAGCGTGG 1259
 QY 1201 CTCTGAGAGATGACGGCCATCTCAGCTGTCTCTGGAGAGAGAGAGGGGGAGAG 1260
 DB 1260 CTCTCTCAGAGATACGGCCATCTCAGCTGTCTCTGGAGAGAGAGGGGGAGAG 1313
 QY 1261 ACAGCCAGCTGAGATCTCTTAAGAGCAATCTTCCGCCCAAGTTCTC- 1319
 DB 1314 ACAGCCAGCTGAGATCTCTTAAGAGCAATCTTCCGCCCAAGTTCTCT 1373
 QY 1320 ----- 1319
 DB 1374 ATCCCTCCAGCACTGAGAGCTGTGCGGAGGTGAGAGAGCTCTTATAGAGAGCTGTGG 1433
 QY 1320 ----- 1319
 DB 1434 TACCAAGGGGCTATCCCTGTGGAGAGATGAGCTGAGCTGTAAAGCACTGGGGACTTC 1493
 QY 1320 ----- 1319
 DB 1494 CTGGTTGGGAGAGCCAGGGAGAGAGATATGATCTGTGTATGTGGATGGCCAC 1553
 QY 1320 -----GAACTGTACCCAGCTGAGAGGGAGGCTTT 1350
 DB 1554 CCCCAGATTTCATCATCCAGTCTTATGATATCTGTATACCTGAGAGGGAGTGGCTTC 1613
 QY 1351 CCTAGCATTTCTTTGCTCATGACCACTTCTGAGAGAGAGAGAGAGAGAG 1410
 DB 1614 CCGAGATACCTCTGCTCATCTCACTGCTGTCTCTCCAGAGAGAGAGAGAG 1673
 QY 1411 AGTGTGTGTCTCTGAGAGAGGGGCTGTGAGAGAGAGAGAGAGAGAGAGAG 1470
 DB 1674 AGTGTGTGTCTCTGAGAGAGGGGCTGTGAGAGAGAGAGAGAGAGAGAG 1733
 QY 1471 CTGGTTGGGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1530
 DB 1734 CTGGTTGGGAGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1793
 QY 1531 CAGAGCCAG 1590
 DB 1794 CAG 1853
 QY 1591 AAGGCCAAGTTTCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1650
 DB 1854 AAGGCCAAGTTTCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1913
 QY 1651 CGTCTCATTTGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1710
 DB 1914 CGTCTCATTTGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1973
 QY 1711 GGGGGGAGACTTCTGAGCTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAG 1770
 DB 1974 GGGGGGAGACTTCTGAGCTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAG 2033
 QY 1771 CTGAGATGAG 1830
 DB 2034 CTGAGATGAG 2093
 QY 1831 CACCGGAGAGCTGTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1890
 DB 2094 CACCGGAGAGCTGTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2153
 QY 1891 GACTTTGGATGTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1950
 DB 2154 GACTTTGGATGTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2213

QY 1951 CAAGTCCCGTGAAGTGAACCGACCTGAGAGCCCTTAACTAGAGGCGCTTACTCTCCGAA 2010
 DB 2214 CAAGTCCCGTGAAGTGAACCGACCTGAGAGCCCTTAACTAGAGAGCTACTCTCAGAG 2273
 QY 2011 AGCAGAGTGTGAGAGCTTTGGCATCTTGTCTGTGGAGAGACTTCAAGCTGTGGGCGCTCC 2070
 DB 2274 AGTGAATGTGTGAGAGCTTTGGCATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2333
 QY 2071 TATCCCAACCTGAG 2130
 DB 2334 TATCCCAACCTGAG 2393
 QY 2131 TCCCAAGAGCTGT 2190
 DB 2394 TCCCAAGAGCTGT 2453
 QY 2191 CCTGGGAG 2250
 DB 2454 CCTGGGAG 2513
 QY 2251 CATCGG 2256
 DB 2514 CATCGG 2519
 RESULT 8
 ID AANT70060 standard; cDNA, 856 BP.
 XX
 AC AANT70060;
 XX
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 27-JAN-1991 (first entry)
 XX
 XX
 DE fes/fps proto-oncogene-related DNA.
 XX
 KM Furin; fur gene; tumour diagnosis; ss.
 XX
 OS Homo sapiens.
 OS Felis catus.
 OS
 FH Key Location/Qualifiers
 FT mat_peptide 3..636
 FT /*tag= a
 FT /product= "fes/fps proto-oncogene protein"
 XX
 PN BP246709-A.
 XX
 PD 25-NOV-1987.
 XX
 PF 19-MAY-1987; 87EP-00200940.
 XX
 PR 20-MAY-1986; 86NL-00001271.
 XX
 PA (UYKA-) UNIV STICHTING KATHOLIEKE.
 XX
 PI Vandeven WJM, Roebroek AJM, Schalcken JA.
 XX
 DR WPI; 1987-328946/47.
 DR P-PSDB; AAP70055.
 XX
 PT Recombinant DNA contg. the fur gene - used for producing furin protein
 PT and antibodies and as a diagnostic aid in the detection of tumours.
 XX
 PS Disclosure; Fig 3; 24pp; English.
 CC The fur gene, encoding furin, is in the human and cat genomes directly
 CC upstream of this fes/fps proto-oncogene sequence. Furin is strongly
 CC expressed in specific types of tumours and labelled RNA or DNA probes of
 CC the fur gene and antibodies against furin can be used for diagnostic
 CC purposes. (See also AANT70061-62 and AAP70056). (Updated on 25-MAR-2003 to
 CC correct PA field.) (Updated on 25-MAR-2003 to correct PI field.) (Updated

CC on 27-AUG-2003 to correct OS field.)
XX Sequence 856 BP; 164 A; 265 C; 259 G; 168 T; 0 U; 0 Other;
SQ
Query Match 28.1%; Score 635; DB 1; Length 856;
Best Local Similarity 100.0%; Pred. No. 2.8e-110;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1622 TGAAGCACTGACAGCCCAACCACTGCTGCTCATTTGCTGTGACCCAGAGAGC 1681
DB 1 TGAAGCACTGACAGCCCAACCACTGCTGCTCATTTGCTGTGACCCAGAGAGC 60
QY 1682 CCATCTACATGCTGATGAGCTTGTGAGGGGGGCGACTTCGACCTTCTCCGACCG 1741
DB 61 CCATCTACATGCTGATGAGCTTGTGAGGGGGGCGACTTCGACCTTCTCCGACCG 120
QY 1742 AGGGGGCCGCTGCGGGGTGAAGACTCTGTGAGATGTGTGGGGGATGACGTGTGCA 1801
DB 121 AGGGGGCCGCTGCGGGGTGAAGACTCTGTGAGATGTGTGGGGGATGACGTGTGCA 180
QY 1802 TGAAGTACTGAGAGCAAGTGTCTGATCCACCGGACCTTGTCTCGGAAGTGCCTGG 1861
DB 181 TGAAGTACTGAGAGCAAGTGTCTGATCCACCGGACCTTGTCTCGGAAGTGCCTGG 240
QY 1862 TGAAGAGAAATGCTGGAAGATGATGATCTTTGGGAGTCCGAGAGAGAGCCGATG 1921
DB 241 TGAAGAGAAATGCTGGAAGATGATGATCTTTGGGAGTCCGAGAGAGAGCCGATG 300
QY 1922 GGGTCTATGAGCCTGAGGGGGCTCAGACAGATCCCGTGAATGGACCGACCTGAG 1981
DB 301 GGGTCTATGAGCCTGAGGGGGCTCAGACAGATCCCGTGAATGGACCGACCTGAG 360
QY 1982 CCCTTAACTAGCGCCGCTACTCTCCGAAAGCAAGTGTGAGCTTTGGCATTTGCTCT 2041
DB 361 CCCTTAACTAGCGCCGCTACTCTCCGAAAGCAAGTGTGAGCTTTGGCATTTGCTCT 420
QY 2042 GGGAGACCTTACGCTGAGGGGGCTCCCGCTATCCCAACCTCAGACATCAGACAGACGG 2101
DB 421 GGGAGACCTTACGCTGAGGGGGCTCCCGCTATCCCAACCTCAGACATCAGACAGACGG 480
QY 2102 AGTTTGTGAGAGAGGGGGCCGTCTGCGCTGCGCCAGAGCTGTCTCGATGCGGTTC 2161
DB 481 AGTTTGTGAGAGAGGGGGCCGTCTGCGCTGCGCCAGAGCTGTCTCGATGCGGTTC 540
QY 2162 GGGTCTATGAGCAAGTGTGAGGGCTTATGAGCTTGGGAGCGGCCCAAGTTGAGACCATCT 2221
DB 541 GGGTCTATGAGCAAGTGTGAGGGCTTATGAGCTTGGGAGCGGCCCAAGTTGAGACCATCT 600
QY 2222 ACCAGAGCTGAGAGACATCCGAAAGGGGATCGG 2256
DB 601 ACCAGAGCTGAGAGACATCCGAAAGGGGATCGG 635
RESULT 9
ADL71052
ID ADL71052 standard; DNA; 2818 BP.
XX
XX ADL71052;
XX
XX 20-MAY-2004 (first entry)
XX
XX Gene encoding type II collagen expression promoting protein, seq id 45.
XX
XX Osteopathic; antiinflammatory; antirheumatic; antiarthritic;
XX gene therapy; type II collagen; expression; cartilage disease;
XX osteoarthritis; cartilage defect; rheumatoid arthritis; human; gene; ds.
XX
XX Homo sapiens.
XX
XX WO2003087375-A1.
XX
XX 23-OCT-2003.
XX

PF 16-APR-2003; 2003WO-JP004802.
XX
XX 16-APR-2002; 2002JP-00113908.
PR 19-APR-2002; 2002US-0373594P.
XX
XX (ASAH) ASAH KASEI KK.
XX
XX Matsuda A, Honda G, Muramatsu S;
XX
XX WPI; 2003-845331/78.
DR P-PSDB; ADL71053.
XX
XX
XX New purified protein that promotes type II collagen expression, useful
PT for preventing and treating a cartilage disease, e.g. osteoarthritis,
PT cartilage defect, or rheumatoid arthritis.
XX
XX
XX Claim 4; SEQ ID NO 45; 271pp; English.
XX
XX
XX The invention relates to a purified protein (I) that promotes type II
CC collagen expression. Also disclosed is an isolated polynucleotide
CC encoding (I), a recombinant vector comprising the polynucleotide and a
CC gene therapeutic agent comprising the recombinant vector as an active
CC ingredient. The proteins, genes, agents and methods are useful for
CC preventing and treating a cartilage disease, e.g. osteoarthritis,
CC cartilage defect, or rheumatoid arthritis. The current sequence
CC represents a human gene of the invention encoding a protein that promotes
CC type II collagen expression.
XX
SQ Sequence 2818 BP; 909 A; 558 C; 684 G; 667 T; 0 U; 0 Other;
Query Match 16.6%; Score 375; DB 1; Length 2818;
Best Local Similarity 64.8%; Pred. No. 3.4e-61;
Matches 606; Conservative 0; Mismatches 520; Indels 9; Gaps 3;
QY 1316 TCTGCACTGTACCGACTGGAAGGGAGGCTTCTGACATTCCTTGTCTATGAC 1375
DB 1686 TCGATATATCTGTATGATTCGAGGGACCGGGTTTCAAAATTCCTCCAGCTTATGATC 1745
QY 1376 ACCTACTGAGCACCCAGACCCCTCAGCAAGAGAGTGTGTCTTCGACAGGGCTG 1435
DB 1746 ACCACTTCAATCAAGAGCAAGTATCATCAGAGAGCTGGGGTGTCTGCTCAACCCA 1805
QY 1436 TGGCCAAAGACAG---TGGGTCTGAACATGAGAGACCTGCTGTGGGTGAGCAGATTG 1492
DB 1806 TCCCAAGAGATAGAAATGGGTCTCAATCATGAAATGTTTCAATGGAGAAATTA 1865
QY 1493 GACGGGGAACCTTGGCGAAGTGTTCAGCGGACGCTCGAGCGCAACAACCTGATG 1552
DB 1866 GCAAGGGGAATTTTGGTGAAGTGTATAGG---GCACTAAAGATTAACCTCTGTTG 1922
QY 1553 CGGTGAAGTCTTGTGAGAGACGCTCCACCTGACCTCAAGGCCAAGTTTCAAGAGAG 1612
DB 1923 CCATTAAGAGTGCAGAGAGAACCTTCTCAGGAATTAAGATTTTCTACAGAGAG 1982
QY 1613 CGAGATCTGGAAGCACTACAGCCCAACATGCTGTCTCATTTGCTGTGACACC 1672
DB 1983 CCAGAAATCTGAGCAATATGATCATCCCAATTTGTCAAACTGATAGGGGTGTGACAC 2042
QY 1673 AGAAGAGCCCATCTACATGCTCATGAGCTTGTGAGGGGGGCGACTTCCTGACCTTCC 1732
DB 2043 AAGAGACGCTGTCTACATCATTTATGAACTGTCTCCAGGGGGTGAATTTTCTACATTTCC 2102
QY 1733 TCCGACAGAGGGGGCCGCTGCGGGGTGAAGACTGTCTGACAGATGTGGGGATGACG 1792
DB 2103 TGAAGAGAGAGAGAGAGAGCTGAGAGCTGAGAGAGAGAGATTTTCTTGAGAGCTTG 2162
QY 1793 CTGCTGGCATGAGATCTGAGAGAGCAAGTGTCTGATCCACCGGACCTTGTCTGCA 1852
DB 2163 CTGCTGGCATGAGATCTGAGAGAGCAAGTGTCTGATCCACAGGAGCTTGTCTGCA 2222
QY 1853 ACTGCTGTGACAGAGAGAAATGCTCTGAAGTCAAGTCACTTTGGAGATGCTCCGAGAG 1912
DB 2223 ACTGCTGTGAGAGAGAAATTAATTAATCTCTGAATATGATGACTTTGAAATGTCTCGGCAAG 2282

XX 23-OCT-2003.
PD 16-APR-2003; 2003WO-JP004802.
XX 16-APR-2002; 2002JP-00113908.
XX 19-APR-2002; 2002US-0373594P.
XX (ASAH) ASAH KASEI KK.
XX Matsumura A, Honda G, Muramatsu S;
XX WPI: 2003-84531/78.
XX P-PSDB; ADL71055.
XX New purified protein that promotes type II collagen expression, useful
XX for preventing and treating a cartilage disease, e.g. osteoarthritis,
XX cartilage defect, or rheumatoid arthritis.
XX Claim 4; SEQ ID NO 47; 271bp; English.
XX The invention relates to a purified protein (I) that promotes type II
XX collagen expression. Also disclosed is an isolated polynucleotide
XX encoding (I), a recombinant vector comprising the polynucleotide and a
XX gene therapeutic agent comprising the recombinant vector as an active
XX ingredient. The proteins, genes, agents and methods are useful for
XX preventing and treating a cartilage disease, e.g. osteoarthritis,
XX cartilage defect, or rheumatoid arthritis. The current sequence
XX represents a human gene of the invention encoding a protein that promotes
XX type II collagen expression.
XX Sequence 2584 BP; 922 A; 446 C; 565 G; 651 T; 0 U; 0 Other;
XX
XX Query Match 14.6%; Score 328.8; DB 11; Length 2584;
XX Best Local Similarity 62.0%; Pred. No. 1.7e-52;
XX Matches 573; Conservative 0; Mismatches 342; Indels 9; Gaps 3;
XX
QY 1321 AACCTGACCGATGGAAGGGAAGGCTTCTGACATCTTCTGATGACCACTTA 1380
DB 1552 AACATATATGATTCGAGGGGCTGTTTCAAACTTCTCACTTATGATCATCAC 1611
QY 1381 CTGAGCACCAGAGCCCTCACCAGAAAGAGTGTGTCTGACAGGGCTGTGCC 1440
DB 1612 TATACACAAAGAGCTCATCTAGAAATCGAGTGTGTCTGATCTTATTCCT 1671
QY 1441 AAGGACAG--TGCGTCTGAACATGAGACCTGGTGTGGTGAGCATTTGACCG 1497
DB 1672 AAGGACAAAGATGATTCATGATGATGATGATGATGATGATGATGATGATG 1731
QY 1498 GGGAACTTTGGCAAGTGTTCAGCGAGCGCTGCGAGCCGACCAACCTGTGGCGTG 1557
DB 1732 GGAATTTTGTGATGATTAAG--GCACATTAAGATTAACCTTCTGTGTCTGTT 1788
QY 1558 AAGCTTTGAGAGAGCGCTCCACCTGACCTCAAGGCGCAAGTTTACAGAGAGAG 1617
DB 1789 AAAACATGTAAGAGATCTTCTCAGGAATTGAAATTAATTTTACAGAGAGCAAA 1848
QY 1618 ATCTGAGAGATGACCAACCCCAATCGTCTGATGATGATGATGATGATGATG 1677
DB 1849 ATCTCAGAGATATGATATCCCAATATGTAATTAAGAGTTTGACACAAAGA 1908
QY 1678 CAGCCCATCTACATGTCATGATGATGATGATGATGATGATGATGATGATG 1737
DB 1909 CAGCCCTGTCTACATCTTATGAGATGATGATGATGATGATGATGATGATG 1968
QY 1738 ACGGAGGGGGCGCGCGGTGAAGACCTGCTGAGATGATGATGATGATGATGATG 1797
DB 1969 AGGAGAGAGATGATCAAAAATCAAACTGATGATGATGATGATGATGATGATG 2028
QY 1798 GGCATGAGATCTGAGAGAGAGATGATGATGATGATGATGATGATGATGATG 1857
DB 2029 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2088

QY 1858 CTGCTGACAGAGAAAGATGCTGAAAGATCAGTATCTTTGGATGTCGGAGAGAGCC 1917
DB 2089 CTGCTGAGTGAATAATATGTTCTGAAATCAGTATCTTTGAAATGTTCTGTCAGAGAT 2148
QY 1918 GATGGGCTCTATGACAGCTCAGGGGCTCAGACAGATGCTCCGATGATGATGATGATG 1977
DB 2149 GGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2205
QY 1978 GAGGCTTAACTACGCGCGCTATCTCTCCGAAAGCAGTGTGATGATGATGATG 2037
DB 2206 GAAGCTTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2265
QY 2038 CTCTGGAGACCTTACGCTGCGGGGCTCCCTCAATCCCACTGCAATGACGAGACA 2097
DB 2266 CTCTGGAGACCTTACGCTGCGGGGCTCCCTCAATCCCACTGCAATGACGAGACA 2325
QY 2098 CCGGAGTTTGTGAG 2157
DB 2326 AGAGAGCAAGTGAAGAGAGATACCGAGATGTCAGCTCCCAAGACCTGCAAGAGATTT 2385
QY 2158 TTCAGCTCATGAG 2217
DB 2386 TCCAAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2445
QY 2218 ATCTACAGAGAGCTGAGAGCATC 2241
DB 2446 CTTGAGAAAGAGCTCATGATCATC 2469

RESULT 12

ACH14291
ID ACH14291 standard; cDNA; 449 BP.
XX ACH14291;
XX 13-OCT-2003 (first entry)
XX
XX Human adult brain cDNA #1503.
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.
XX Homo sapiens.
XX US2003073623-A1.
XX 17-APR-2003.
XX 30-JUL-2001; 2001US-00918995.
XX 30-JUL-2001; 2001US-00918995.
XX 30-JUL-2001; 2001US-00918995.
XX (DRMA/) DRMANAC R T.
XX (LABA/) LABAT I.
XX (STAC/) STACHE-CRAIN B.
XX (DICK/) DICKSON M C.
XX (JONES/) JONES L W.
XX Dermanac RT, Labat I, Steche-Crain B, Dickson MC, Jones LW;
XX WPI: 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.
XX Claim 1; SEQ ID NO 1503; 449bp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to

CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030073623
CC
XX
SQ Sequence 449 BP; 90 A; 124 C; 148 G; 83 T; 0 U; 4 Other;

Query Match 14.5%; Score 326.8; DB 9; Length 449;
Best Local Similarity 99.4%; Pred. No. 3.3e-52;
Matches 328; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1707 GCAGGGGGGGGAGCTTCTGACCTTCCGCGCAGAGAGGGGGCCGCTGCGGGTGAAGAC 1766
DB 120 GCAGGGGGGGGAGCTTCTGACCTTCCGCGCAGAGAGGGGGCCGCTGCGGGTGAAGAC 179
QY 1767 TCTGCTGAGATGATGGGGGATCAGCTGCTGCGCATGAGTACCTGAGAGCAAGTCTG 1826
DB 180 TCTGCTGAGATGATGGGGGATCAGCTGCTGCGCATGAGTACCTGAGAGCAAGTCTG 239
QY 1827 CATCCACCGGAGCTGCTGCTGCGCATGCTGCTGCGCATGAGAGAGATCTCTGAAGAT 1886
DB 240 CATCCACCGGAGCTGCTGCTGCGCATGCTGCTGCGCATGAGAGAGATCTCTGAAGAT 299
QY 1887 CATGACTTTGGGATGTCGCGAGAGAGAGCGGATGATGAGAGCTTCAAGGGGCGCT 1946
DB 300 CATGACTTTGGGATGTCGCGAGAGAGAGCGGATGATGAGAGCTTCAAGGGGCGCT 359
QY 1947 CAGACAGATCCCGGTGAAGTGAACCGACCTGAGCCCTTAAGTACGCGCTACTCTTC 2006
DB 360 CAGACAGATCCCGGTGAAGTGAACCGACCTGAGCCCTTAAGTACGCGCTACTCTTC 419
QY 2007 CGAAGCGAGCTGTGAGCTTTGGCATCTT 2036
DB 420 CGAAGCGAGCTGTGAGCTTTGGCATCTT 449

RESULT 13
AAS17052
ID AAS17052 standard; cDNA; 361 BP.
XX
AC AAS17052;
XX
DT 14-FEB-2002 (first entry)
XX
DE Human protein kinase cDNA mine16676human_sl.
XX
XX Human; 88; protein kinase; mine16676human_sl; cytosolic; antiangiinal;
XX hypotensive; cardiac; cardiovascular disorder; heart failure;
XX hypertensive; arterial fibrillation; dilated cardiomyopathy;
XX idiopathic cardiomyopathy; angina; proliferative disorder; cancer;
XX melanoma; prostate cancer; cervical cancer; breast cancer; colon sarcoma;
XX gene therapy.
XX
OS Homo sapiens.
XX
XX US6309849-B1.
XX
XX 30-OCT-2001.
XX
XX 31-AUG-1999; 99US-00387212.
XX
XX 31-AUG-1999; 99US-00387212.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
PA

XX
PI Robison KE;
XX
XX WPI; 2002-048371/06.
XX
DR
XX
XX Identifying compound which binds to a kinase, useful for treating
XX diseases e.g. cancer, by contacting kinase with test compound and
XX detecting its binding to the kinase.
XX
XX
XX Claim 1; Fig 5; 45pp; English.
XX
XX
XX The invention relates to identifying a compound which binds to a kinase
XX (encoded by a nucleotide sequence of 1868, 403, 545, 361, 473, 3001, 526,
XX 683 or 1448 base pair (bp) as given in the specification) comprising
XX contacting the kinase with a test compound under suitable conditions for
XX binding, and detecting binding of the compound to the kinase. The method
XX is useful for identifying a compound which binds to the kinase and also
XX for isolating compounds which modify the activity of the kinase. The
XX identified compounds are useful for treating a subject having a disorder
XX characterised by aberrant kinase activity where the disorder includes
XX cellular growth related disorders which includes a disorder, disease, or
XX condition characterised by a deregulation, e.g. an upregulation or a
XX downregulation, of cellular growth, cellular growth deregulation due to
XX deregulation of cellular proliferation, cell cycle progression, cellular
XX differentiation and/or cellular hypertrophy, cardiovascular disorders
XX such as heart failure, hypertension, arterial fibrillation, dilated
XX cardiomyopathy, idiopathic cardiomyopathy, or anginal, proliferative
XX disorders such as cancer (including melanoma, prostate cancer, cervical,
XX breast, colon sarcoma). The kinases and antibodies raised against them
XX are useful in one or more method such as screening assays, predictive
XX medicine and methods of treatment. The nucleic acid molecules are useful
XX for expressing kinase and phosphatase protein (e.g. in gene therapy
XX applications), to detect kinase and phosphatase mRNA or a genetic
XX alteration in a kinase and phosphatase gene and to modulate kinase and
XX phosphatase activity. The present sequence is a cDNA for a human protein
XX kinase
XX
SQ Sequence 361 BP; 71 A; 101 C; 120 G; 68 T; 0 U; 1 Other;

Query Match 13.1%; Score 296.2; DB 6; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.9e-46;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

QY 1613 CGAGATCTTGAAGCAGTACAGACCCCAACATCGTGGCTCATTTGTTGACCC 1672
DB 20 CNAAGATCTTGAAGCAGTACAGACCCCAACATCGTGGCTCATTTGTTGACCC 79
QY 1673 AGAAGCAGCCCATCTACATGCTGATGAGCTTGTGAGAGGGGGGCACTTCTACCTTC 1732
DB 80 AGAAGCAGCCCATCTACATGCTGATGAGCTTGTGAGAGGGGGGCACTTCTACCTTC 139
QY 1733 TCCGCAAGGAGGGGGGCGCGCTGCGGGGTGAAGACTGTGCTGAGATGTTGGGATGACG 1792
DB 140 TCCGCAAGGAGGGGGGCGCGCTGCGGGGTGAAGACTGTGCTGAGATGTTGGGATGACG 199
QY 1793 CTGCTGCGATGAGTACTCTGAGAGCAAGTGTCTGATCCACCGGACCTTGTCTCGA 1852
DB 200 CTGCTGCGATGAGTACTCTGAGAGCAAGTGTCTGATCCACCGGACCTTGTCTCGA 259
QY 1853 ACTGCTGTGAGAGAGAGATGTTCTGAG-ATCAGTACCTTT-GGATGTTCCGAGA 1910
DB 260 ACTGCTGTGAGAGAGAGATGTTCTGAGAGATGATGATGATGTTGGGGATGTTCCGAGA 319
QY 1911 GGAAGCC---GATGGGGTCTATGACAGCTCAGGGGGGCTTAC 1949
DB 320 GGAAGCCGATTTGGGGGTCTATGACAGCTCAGGGGGGCTTAC 361

RESULT 14
AAD61604
ID AAD61604 standard; cDNA; 361 BP.
XX
XX AAD61604;
XX
AC

```
XX
DE 15-JUN-2004 (first entry)
XX
DE Human protein kinase CDNA, 16676S1.
XX
KM Kinase; phosphatase; drug target; therapy; gene; human; ss.
XX
OS Homo sapiens.
XX
PN US2003104505-A1.
XX
PD 05-JUN-2003.
XX
PF 12-APR-2002; 2002US-00121925.
XX
PR 31-AUG-1999; 99US-00387212.
XX
PR 07-SEP-2001; 2001US-00948802.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Robison KE;
XX
DR WPI; 2003-801234/75.
XX
PT Detecting and modulating the activity of specified kinases and
XX phosphatases which are potentially useful as drug targets.
XX
PS Claim 1; Fig 5; Opp; English.
XX
XX The present invention relates to a method for detecting the presence of a
XX kinases or phosphatases encoded by nucleotides. The kinases and
XX phosphatases and their encoding nucleic acids are potentially useful as
XX drug targets. The present invention may also be useful in diagnosing
XX disease. The present sequence is human protein kinase CDNA
XX
SQ Sequence 361 BP; 71 A; 101 C; 120 G; 68 T; 0 U; 1 Other;
XX
Query Match 13.1%; Score 296.2; DB 10; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.9e-46;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;
XX
QY 1613 CGAGATCTCGAAGCAGTACAGCCACCCCAACATCGGCTCATTTGATGTCGACCC 1672
DB 20 CNAAGATCTCGAAGCAGTACAGCCACCCCAACATCGGCTCATTTGATGTCGACCC 79
QY 1673 AGAAGCAGCCCATCTACATGCTGATGAGCTTGTGAGGGGGCGACTTCTGACCTTCC 1732
DB 80 AGAAGCAGCCCATCTACATGCTGATGAGCTTGTGAGGGGGCGACTTCTGACCTTCC 139
QY 1733 TCCGACGAGAGGGGGCCCGCTCGGGGTGAAGACTCTGCTGCAGATGTGGGGATGCG 1792
DB 140 TCCGACGAGAGGGGGCCCGCTCGGGGTGAAGACTCTGCTGCAGATGTGGGGATGCG 199
QY 1793 CTGCTGGCATGTGAGTACTCTGAGAGCAAGTCTGACCCCGGGACCTGGCTGCTCGGA 1852
DB 200 CTGCTGGCATGTGAGTACTCTGAGAGCAAGTCTGACCCCGGGACCTGGCTGCTCGGA 259
QY 1853 ACTGCTGTGACAGAGAAATGCTCTGAAG-ATCAGTACTTT-GGAGTGTCCGAGA 1910
DB 260 ACTGCTGTGACAGAGAAATGCTCTGAAGAAATCAGTACTTTGGGAGATGCCGAGA 319
QY 1911 GGAAGCC--GATGGGCTTATGCAAGCTTCAGGGGGGCTTCAG 1949
DB 320 GGAAGCCCGATTTGGGGTCTATGCAAGCTTCAGGGGGGCTTCAG 361
XX
RESULT 15
ABX14976
ID ABX14976 standard; cDNA; 361 BP.
XX
AC ABX14976;
XX
DT 13-MAR-2003 (first entry)
```

```
XX
DE Human protein kinase CDNA mine16676human_sl.
XX
DE Human; ss; protein kinase; mine16676human_sl; cancer; cytostatic;
XX anti-anginal; hypotensive; cardiact. proliferative disorder;
XX cellular growth related disorder; cardiovascular disorder; heart failure;
XX hypertension; atrial fibrillation; dilated cardiomyopathy;
XX idiopathic cardiomyopathy; angina.
XX
OS Homo sapiens.
XX
PN US6465232-B1.
XX
PD 15-OCT-2002.
XX
PF 07-SEP-2001; 2001US-00948802.
XX
PR 31-AUG-1999; 99US-00387212.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Robison KE;
XX
DR WPI; 2003-147067/14.
XX
PT Novel human kinase and phosphatase nucleic acid molecules useful for
XX treating cellular proliferative disorders such as cancer, cardiovascular
XX diseases, hypertension, heart failure and angina.
XX
PS Claim 7; Fig 5; 47pp; English.
XX
XX The invention relates to an isolated human kinase and phosphatase nucleic
XX acid molecule appearing as ABX14972, ABX14973, ABX14975, ABX14976, and
XX ABX14978-ABX14981, or their complement. Also included are: (1) an
XX isolated nucleic acid molecule which is at least 90 % identical to the
XX nucleotide sequence ABX14973 or ABX14975, or 95 % identical to the
XX complement, where the nucleic acid molecule encodes a polypeptide having
XX a kinase activity; (2) An isolated nucleic acid molecule which hybridises
XX to ABX14973, ABX14975, ABX14976, and ABX14978 in 6X saline sodium citrate
XX (SSC) at 45 plusoc, followed by one or more washes in 0.2X SSC, 0.1%
XX sodium dodecyl sulphate (SDS) at 65 plusoc, where the molecule encodes a
XX polypeptide with kinase activity; (3) a vector comprising the
XX polynucleotides, operatively linked to a recombinant regulatory sequence;
XX and (4) expressing a polypeptide by culturing a host cell comprising the
XX vector under conditions in which the nucleic acid molecule is expressed.
XX The nucleic acids are useful as modulating agents in regulating a variety
XX of cellular processes, and fragments are useful as primers or
XX hybridisation probes for detecting kinase and phosphatase encoding
XX nucleic acids. The nucleic acids are useful for treating proliferative
XX disorders such as cancer and cellular growth related disorders including
XX cardiovascular disorders such as heart failure, hypertension, atrial
XX fibrillation, dilated cardiomyopathy, idiopathic cardiomyopathy or
XX angina. The present sequence is one of the human kinase or phosphatase
XX cDNAs of the invention
XX
SQ Sequence 361 BP; 71 A; 101 C; 120 G; 68 T; 0 U; 1 Other;
XX
Query Match 13.1%; Score 296.2; DB 10; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.9e-46;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;
XX
QY 1613 CGAGATCTCGAAGCAGTACAGCCACCCCAACATCGGCTCATTTGATGTCGACCC 1672
DB 20 CNAAGATCTCGAAGCAGTACAGCCACCCCAACATCGGCTCATTTGATGTCGACCC 79
QY 1673 AGAAGCAGCCCATCTACATGCTGATGAGCTTGTGAGGGGGCGACTTCTGACCTTCC 1732
DB 80 AGAAGCAGCCCATCTACATGCTGATGAGCTTGTGAGGGGGCGACTTCTGACCTTCC 139
QY 1733 TCCGACGAGAGGGGGCCCGCTCGGGGTGAAGACTCTGCTGCAGATGTGGGGATGCG 1792
DB 140 TCCGACGAGAGGGGGCCCGCTCGGGGTGAAGACTCTGCTGCAGATGTGGGGATGCG 199
```

```

QY      1793 CTGCTGGCATGAGTACCTGAGAGCAAGTGTGCATCCACCGGACCTGGTGTGCGA 1852
Db      200 CTGCTGGCATGAGTACCTGAGAGCAAGTGTGCATCCACCGGACCTGGTGTGCGA 259
QY      1853 ACTGCTGTGTGACAGAGAGAAATGTCTGAG-ATCAGTACTTT-GGATGTCCGAGA 1910
Db      260 ACTGCTGTGTGACAGAGAAATGTCTGAGAAATCAGTGTGCTTTGGGATGTCCGAGA 319
QY      1911 GGAAGCC--GATGGGCTTATGACAGCTCAGGGGGCTCAG 1949
Db      320 GGAAGCCCGATTGGGGCTTATGACAGCTCAGGGGGCTCAG 361

```

Search completed: March 19, 2005, 04:31:26
 Job time : 1087.29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 03:18:16 ; Search time 9383.68 Seconds
(without alignments)
11649.479 Million cell updates/sec

Title: US-10-660-763-1_COPY_72_2327

Perfect score: 2256

Sequence: 1 atggcgtctctctctgcgact.....gcattccgaagcgcgcgcgcg 2256

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank1:
1: gb_ba:*
2: gb_ncg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_xo:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2256	100.0	2674	6	AR183261	AR183261 Sequence
2	2256	100.0	2674	6	AR456323	AR456323 Sequence
3	2036	90.2	2697	6	CQ725220	CQ725220 Sequence
4	2036	90.2	2788	9	BC035357	BC035357 Homo sapi
5	2029.6	90.0	2771	9	HSCFES	X52192 H. sapiens R
6	2023.4	89.7	2889	6	AX281847	AX281847 Sequence
7	1540.4	68.3	2680	10	WMPESCR	X12616 Mouse c-fes
8	1018.2	45.1	2715	14	ACFTS140A	M14930 Fujinami sa
9	1018.2	45.1	4788	14	ACF	J03194 Fujinami sa
10	1018.2	45.1	4788	14	AF033810	AF033810 Fujinami
11	841.4	37.3	4901	13	M22820	M22820 Figure 4. N
12	821.8	36.4	2397	14	PCGSTONC	J02088 Feline sarc
13	810.8	35.9	2946	14	PCGGAONC	J02087 feline sarc
14	656.4	29.1	4788	5	BC073445	BC073445 Xenopus l
15	635	28.1	856	6	A06935	A06935 feline sarc
16	626.2	27.8	2284	14	ACSGAGFPS	K01690 Avian sarco
17	443	19.6	772	10	BC038130	BC038130 Mus muscu
18	376.6	16.7	1563	10	BC051249	BC051249 Mus muscu
19	376.6	16.7	2948	10	BC058100	BC058100 Mus muscu

20	376.6	16.7	2994	10	MMU76762	U76762 Mus musculi
21	375	16.6	2069	10	MUSPBR	M32054 Mouse tyros
22	374.8	16.6	1171	10	RNPLK	X13412 Rat mRNA fo
23	342.8	15.2	2947	4	AF187884	AF187884 Canis fam
24	328.8	14.6	2950	9	HUMTKFER	J03358 Human tyros
25	318.4	14.1	3882	3	SREPSTK	Y17051 Sycon rapha
26	296.2	13.1	361	6	AR175607	AR175607 Sequence
27	296.2	13.1	361	6	AR236724	AR236724 Sequence
28	274.4	12.2	2786	3	AB006567	AB006567 Ephratia
29	271.6	12.0	3198	6	CQ597087	CQ597087 Sequence
30	271.6	12.0	3369	6	CQ591168	CQ591168 Sequence
31	271.6	11.9	4972	3	BT003462	BT003462 Drosophila
32	268.4	11.9	2303	3	DMU50450	U50450 Drosophila
33	266.8	11.8	3297	3	DMPS85D	X52844 Drosophila
34	230.6	10.2	2597	6	CQ726604	CQ726604 Sequence
35	227.2	10.1	15297	6	AR183262	AR183262 Sequence
36	227.2	10.1	15297	6	AR456324	AR456324 Sequence
37	225.6	10.0	12263	9	HSPESFPS	X06292 Human c-fes
38	224.6	10.0	142201	2	AC124248	AC124248 Homo sapi
39	224.6	10.0	156889	2	AC004586	AC004586 Homo sapi
40	211.4	9.4	4160	6	A06936	A06936 H. sapiens f
41	211.4	9.4	4160	6	E01406	E01406 DNA sequenc
42	201.2	8.9	2398	9	HSRTKSPH	Z27409 H. sapiens m
43	198	8.8	2955	6	AX481387	AX481387 Sequence
44	198	8.8	3370	6	CQ776660	CQ776660 Sequence
45	198	8.8	3370	6	ARS59315	ARS59315 Sequence

ALIGNMENTS

RESULT 1
AR183261
LOCUS AR183261 2674 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 1 from patent US 6340584.
ACCESSION AR183261
VERSION AR183261.1 GI:20226854
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2674)
Gan, W., Ye, J., Di Francesco, V. and Beasley, E. M.
Isolated human kinase proteins, and uses thereof
human Kinase proteins, and uses thereof
Patent: US 6340584-A 1 22-JAN-2002;
JOURNAL Location/Qualifiers
FEATURES
source
1..2674
/organism="Unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 2256; DB 6; Length 2674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCTCTCTCTTCTGAGCTGTGACAGCCCGGACGCGGAGTGCAGCAATGCG 60
DB 72 ATGGGCTCTCTCTTCTGAGCTGTGACAGCCCGGACGCGGAGTGCAGCAATGCG 121
QY 61 GAGGCGGACCTTCCTTACTGAGAGGCGGAGGAGGAGTGGCCGAGGGGCTGAGAGT 120
DB 132 GAGGCGGACCTTCCTTACTGAGAGGCGGAGGAGGAGTGGCCGAGGGGCTGAGAGT 191
QY 121 GACAGGAGTATGACAGAGTCTTCAACATGTCCTGACAGAGTGGGGGCGACAGC 180
DB 192 GACAGGAGTATGACAGAGTCTTCAACATGTCCTGACAGAGTGGGGGCGACAGC 251
QY 181 CGAGGCGATGACCGCTTACAGCCCGCATGCTGAGTGGGCTGAGATCAACAGCCAACT 240
DB 252 CGGCGCATGACCGCTTACAGCCCGCATGCTGAGTGGGCTGAGATCAACAGCCAACT 311
QY 241 GAGGCGCTGAGCGCTGCTGTGGGCGACGACGAGGAGTCTGAACCTCAGGGCCCTGAGC 300

Db 312 GAGGGCTGAGCGCTTGTGGGGGAGCAAGAGAGATCTGAATCAAGGGCCCTCGAGC 371
Qy 301 AAGCTGAGCTGCTTATCCGGGAACGGCAGAGCTTTCGAAAGCTTACAGCGAGCTGG 360
Db 372 AAGCTGAGCTGCTTATCCGGGAACGGCAGAGCTTTCGAAAGCTTACAGCGAGCTGG 431
Qy 361 CAGCAGCTGAGAGAGAGCTCAACAAAGCCACAGCGAGGACTTGAAGAGCTGAAGAGC 420
Db 432 CAGCAGCTGAGAGAGAGCTCAACAAAGCCACAGCGAGGACTTGAAGAGCTGAAGAGC 491
Qy 421 CAGTACGAGCTTGGCAAGGACAGTGGCCAAAGCCAAAGGCAAGTACCAAGAGGCGCAGC 480
Db 492 CAGTACGAGCTTGGCAAGGACAGTGGCCAAAGCCAAAGGCAAGTACCAAGAGGCGCAGC 551
Qy 481 AAGAGCAAGAGCCGTGAACAAGGCCAAGCAAGTATGTGGCAGCCTGTGGAAGCTCTTT 540
Db 552 AAGAGCAAGAGCCGTGAACAAGGCCAAGCAAGTATGTGGCAGCCTGTGGAAGCTCTTT 611
Qy 541 GCTCACCAACAGCTATGTGCTGGGGGTGGGGGCTGGCAGCTTACACCAAGCAGCAGC 600
Db 612 GCTCACCAACAGCTATGTGCTGGGGGTGGGGGCTGGCAGCTTACACCAAGCAGCAGC 671
Qy 601 CACCAAGCTCTGTGCTGGGGGTGGGGGCTGTGCGGTCACTGCAAGGACCTGCAAGAGATGGCT 660
Db 672 CACCAAGCTCTGTGCTGGGGGTGGGGGCTGTGCGGTCACTGCAAGGACCTGCAAGAGATGGCT 731
Qy 661 TGCATCTTGAAGAGATCTCTGAGAGAAATCTGGAATTAAGCAGCTGTGGAGATGAG 720
Db 732 TGCATCTTGAAGAGATCTCTGAGAGAAATCTGGAATTAAGCAGCTGTGGAGATGAG 791
Qy 721 GTGGTGGCCATTCACCGGAGAGATGGCTGAGCTGTGCGGCACTGCAAGCCTGAGGCTGAG 780
Db 792 GTGGTGGCCATTCACCGGAGAGATGGCTGAGCTGTGCGGCACTGCAAGCCTGAGGCTGAG 851
Qy 781 TACCAAGGCTTCTGCGAAGATGGGGTCCGACCTGACGTCCACCTGTGTCACTTTC 840
Db 852 TACCAAGGCTTCTGCGAAGATGGGGTCCGACCTGACGTCCACCTGTGTCACTTTC 911
Qy 841 GATGATCTACTGCTTGAAGAGAGGTGAACCGCTGGAAGCTTGGGAGAGCTTCCAGCTGAACGAG 900
Db 912 GATGATCTACTGCTTGAAGAGAGGTGAACCGCTGGAAGCTTGGGAGAGCTTCCAGCTGAACGAG 971
Qy 901 CTGACTGTGAGAGCGTGCAGCAACGCTGACCTGAGTGAAGATGAGCTGTGGTGGCC 960
Db 972 CTGACTGTGAGAGCGTGCAGCAACGCTGACCTGAGTGAAGATGAGCTGTGGTGGCC 1031
Qy 961 ACCGAGATGCTGTTCAAGCGGAGAGAGATGGTTACGCAAGCTGCAACAGAGCTCCGGAAT 1020
Db 1032 ACCGAGATGCTGTTCAAGCGGAGAGAGATGGTTACGCAAGCTGCAACAGAGCTCCGGAAT 1091
Qy 1021 GAAGAGAGAAACCAACCCCGGAGAGCGGAGTCACTGTGGGCAAGAGCAAGTGTCTG 1080
Db 1092 GAAGAGAGAAACCAACCCCGGAGAGCGGAGTCACTGTGGGCAAGAGCAAGTGTCTG 1151
Qy 1081 CAAGAAGCACTGAGAGGGGCTGCAGGTAGCTGTGCAGCGGCAAGCTGAGGCGCCAG 1140
Db 1152 CAAGAAGCACTGAGAGGGGCTGCAGGTAGCTGTGCAGCGGCAAGCTGAGGCGCCAG 1211
Qy 1141 CAGAGATGCTGCAGAACCAAGCTGAGCACTTGGGCGGCGAGGCGGCGCTGTGTCTG 1200
Db 1212 CAGAGATGCTGCAGAACCAAGCTGAGCACTTGGGCGGCGAGGCGGCGCTGTGTCTG 1271
Qy 1201 CTCTCTGAGATACCGGCACTCCAGTGTCTCTCGGAGAGAGAGAGAGAGGGGGGAAG 1260
Db 1272 CTCTCTGAGATACCGGCACTCCAGTGTCTCTCGGAGAGAGAGAGAGGGGGGAAG 1331
Qy 1261 ACACCAAGCTGAGATCTTAAAGAGCAATCTCAAGAAATCTTCCGCGCCCAAGTCTCG 1320
Db 1332 ACACCAAGCTGAGATCTTAAAGAGCAATCTCAAGAAATCTTCCGCGCCCAAGTCTCG 1391
Qy 1321 AACCTGTACGAGTGAAGGGGAAGGCTTCTAGCAATTCCTTGTCTCATGACCACTA 1380

Db 1392 AACCTGTACGAGTGAAGGGGAAGGCTTCTAGCAATTCCTTGTCTCATGACCACTA 1451
Qy 1381 CTGAGACCCAGCAGCCCTCTCAAGAAAGATGGTGTCTGCAAGAGGCTGTGCC 1440
Db 1452 CTGAGACCCAGCAGCCCTCTCAAGAAAGATGGTGTCTGCAAGAGGCTGTGCC 1511
Qy 1441 AAGCAAGTGGGTCTGAACCATGAGGACTGTGTGGGTGAGCAGATTTGAAGCGGGG 1500
Db 1512 AAGCAAGTGGGTCTGAACCATGAGGACTGTGTGGGTGAGCAGATTTGAAGCGGGG 1571
Qy 1501 AACTTGGGGAAGTGTTCAGCGGAGCGCTTGCAGAGCCGCAACAACCTGTGTGGCTGAAG 1560
Db 1572 AACTTGGGGAAGTGTTCAGCGGAGCGCTTGCAGAGCCGCAACAACCTGTGTGGCTGAAG 1631
Qy 1561 TCTTGTGAGAGAGCGTCCCACTGACCTCAAGGGCAAGTTTCTACAGGAAGAGGATC 1620
Db 1632 TCTTGTGAGAGAGCGTCCCACTGACCTCAAGGGCAAGTTTCTACAGGAAGAGGATC 1691
Qy 1621 CTGAAGCAGTACAGCCACCCCAACATCGTCTCTCATTTGTGTGCAACCAAGAGCAG 1680
Db 1692 CTGAAGCAGTACAGCCACCCCAACATCGTCTCTCATTTGTGTGCAACCAAGAGCAG 1751
Qy 1681 CCCATCTACATGTGTCAAGAGCTTGTGCAGGGGGGCGACTTCTGACCTTCCGCAAG 1740
Db 1752 CCCATCTACATGTGTCAAGAGCTTGTGCAGGGGGGCGACTTCTGACCTTCCGCAAG 1811
Qy 1741 GAGGGGGGCGCGCTGCGGGGTGAAGACTGTGCTGAGATGGTGGGGGATGACAGCTGTGCGC 1800
Db 1812 GAGGGGGGCGCGCTGCGGGGTGAAGACTGTGCTGAGATGGTGGGGGATGACAGCTGTGCGC 1871
Qy 1801 ATGAGTACCTTGAAGAGCAAGTGTGATCAACCGGGAACCTGAGCTGTGGAACCTGCTG 1860
Db 1872 ATGAGTACCTTGAAGAGCAAGTGTGATCAACCGGGAACCTGAGCTGTGGAACCTGCTG 1931
Qy 1861 GTGACAGAGAAAGATGTCTGAAGATCAAGTACTTTGGGATGTCCCGAAGAGAGCCGAT 1920
Db 1932 GTGACAGAGAAAGATGTCTGAAGATCAAGTACTTTGGGATGTCCCGAAGAGAGCCGAT 1991
Qy 1921 GGGGTCTATGAGCGCTCAGGGGGGCTCAGACAAGTCCCGTGAAGTGAACCGCACTGAG 1980
Db 1992 GGGGTCTATGAGCGCTCAGGGGGGCTCAGACAAGTCCCGTGAAGTGAACCGCACTGAG 2051
Qy 1981 GCCCTTAATCAAGCGCGCTACTCTCCGAAGCGAGGTGAGACTTTGGCATTTGCTC 2040
Db 2052 GCCCTTAATCAAGCGCGCTACTCTCCGAAGCGAGGTGAGACTTTGGCATTTGCTC 2111
Qy 2041 TGGAGACCTTCAAGCTTGGGGGCTTCCCTATTCCTCAACTCAGCAATCAGCAGACAGG 2100
Db 2112 TGGAGACCTTCAAGCTTGGGGGCTTCCCTATTCCTCAACTCAGCAATCAGCAGACAGG 2171
Qy 2101 GAGTTTGTGAGAAAGGGGGGCGGTCTGAGCTGCGGAGAGCTGTGTCTGATGCGGTGTC 2160
Db 2172 GAGTTTGTGAGAAAGGGGGGCGGTCTGAGCTGCGGAGAGCTGTGTCTGATGCGGTGTC 2231
Qy 2161 AAGCTCATAGAGCAGTGTGGGCTATGAGCCTTGGGAGCGGCGCCAGCTTCAAGCAATC 2220
Db 2232 AAGCTCATAGAGCAGTGTGGGCTATGAGCCTTGGGAGCGGCGCCAGCTTCAAGCAATC 2291
Qy 2221 TACCAAGAGCTCAGAGCATCCGAAAGCGGCAATCGG 2256
Db 2292 TACCAAGAGCTCAGAGCATCCGAAAGCGGCAATCGG 2327

RESULT 2
AR456323
LOCUS AR456323 2674 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 1 from patent US 6686187.
ACCESSION AR456323
VERSION AR456323.1 GI:42691378
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 2674)
AUTHORS Gan W., Ye J., Di Francesco, V. and Beasley, E.M.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding
human kinase proteins, and uses thereof
JOURNAL Patent: US 6686187-A 1 03-FEB-2004;
FEATURES Location/Qualifiers
source 1..2674
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 2256; DB 6; Length 2674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGGCTTCTCTCTAGCTGTGACGCCCCGAGGCGACGGGGTCTCTGCAGCAATGACG 60
2 ATGGGCTTCTCTCTAGCTGTGACGCCCCGAGGCGACGGGGTCTCTGCAGCAATGACG 131
3 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
4 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 191
5 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
6 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 251
7 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
8 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 252
9 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 311
10 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
11 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 371
12 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
13 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 431
14 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
15 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 491
16 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
17 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 551
18 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
19 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 611
20 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
21 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 671
22 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
23 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 731
24 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
25 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 791
26 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
27 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 851
28 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
29 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 911
30 GAGGCGGAGCTTGTCTAATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900

912 GATGATGCTACTGCTTGAAGAGGAGTGAACCGCTGAGGCTGAGGAGCTCCAGCTGAACGAG 971
901 CTGACTGTGAGAGGAGTGAACCGCTGAGGCTGAGGCTGAGGAGCTGAGGAGCTGAGG 960
972 CTGACTGTGAGAGGAGTGAACCGCTGAGGCTGAGGCTGAGGAGCTGAGGAGCTGAGG 1031
961 ACCGAGATGAGTGAACCGCTGAGGCTGAGGCTGAGGCTGAGGAGCTGAGGAGCTGAGG 1020
1032 ACCGAGATGAGTGAACCGCTGAGGCTGAGGCTGAGGCTGAGGAGCTGAGGAGCTGAGG 1091
1021 GAAAG 1080
1092 GAAAG 1151
1081 GAAAG 1140
1152 GAAAG 1211
1141 GAAAG 1200
1212 GAAAG 1271
1201 GAAAG 1260
1272 GAAAG 1331
1261 GAAAG 1320
1332 GAAAG 1391
1321 GAAAG 1380
1392 GAAAG 1451
1381 GAAAG 1440
1452 GAAAG 1511
1441 GAAAG 1500
1512 GAAAG 1571
1501 GAAAG 1560
1572 GAAAG 1631
1561 GAAAG 1620
1632 GAAAG 1691
1621 GAAAG 1680
1692 GAAAG 1751
1681 GAAAG 1740
1752 GAAAG 1811
1741 GAAAG 1800
1812 GAAAG 1871
1801 GAAAG 1860
1872 GAAAG 1931
1861 GAAAG 1920
1932 GAAAG 1991
1921 GAAAG 1980
1992 GAAAG 2051

QY 1981 GCCCTTAACGAGCGCTACTCTCCGAAAGGAGCTGTGAGCTTTGGCATCTTGCTC 2040
DB 2052 GCCCTTAACGAGCGCTACTCTCTCCGAAAGGAGCTGTGAGCTTTGGCATCTTGCTC 2111
QY 2041 TGGGAGACCTTACGCTGCGGGGCTCCCTCTATCCCACTCAGCAATCAGCAGACCG 2100
DB 2112 TGGGAGACCTTACGCTGCGGGGCTCCCTCTATCCCACTCAGCAATCAGCAGACCG 2171
QY 2101 GAGTTGTGAGAAAGGGGGGCGCTGCTCCCTGCCAGAGCTGTGTCTGATGCGGTTC 2160
DB 2172 GAGTTGTGAGAAAGGGGGGCGCTGCTCCCTGCCAGAGCTGTGTCTGATGCGGTTC 2231
QY 2161 AGGCTCATGAGAGCTGCTGAGCTAGCTAGCTGAGGAGCGGCCAGCTTACAGCATC 2220
DB 2232 AGGCTCATGAGAGCTGCTGAGCTAGCTAGCTGAGGAGCGGCCAGCTTACAGCATC 2291
QY 2221 TACCAAGAGCTGCAAGAGCATCCGAAAGCGGATCG 2256
DB 2292 TACCAAGAGCTGCAAGAGCATCCGAAAGCGGATCG 2327

RESULT 3
LOCUS CQ725220 2697 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 11154 from Patent WO2068579.
ACCESSION CQ725220
VERSION CQ725220.1 GI:42286077

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanelements or transcripts, for detecting expression and other uses
thereof

JOURNAL Patent: WO 02068579-A 11154 06-SEP-2002;
PE Corperation (NY) (US)

FEATURES
Location/Qualifiers

1..2697
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 90.2%; Score 2036; DB 6; Length 2697;
Best Local Similarity 91.5%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 210; Gaps 1;

QY 1 ATGGGCTTCTTCTGAGCTGTGACGCCCCAGGCGGCTCTGCGCAAAATGACG 60
DB 11 ATGGGCTTCTTCTGAGCTGTGACGCCCCAGGCGGCTCTGCGCAAAATGACG 70
QY 61 GAGGCGAGCTTCTGCTACTGAGAGGATGAGAAAGTGGATGGCCAGCGGCTCAAGAGT 120
DB 71 GAGGCGAGCTTCTGCTACTGAGAGGATGAGAAAGTGGATGGCCAGCGGCTCAAGAGT 130
QY 121 GACAGGAGATGACAGAGCTGCTTACACACATGTCCTGACAGACAGTGGGGCCAGAGC 180
DB 131 GACAGGAGATGACAGAGCTGCTTACACACATGTCCTGACAGACAGTGGGGCCAGAGC 190
QY 181 CGGGCCATCAGCCCTGACAGCCCATCATGATGATGCTGGGCTGAGATCAACAGCAAACT 240
DB 191 CGGGCCATCAGCCCTGACAGCCCATCATGATGATGCTGGGCTGAGATCAACAGCAAACT 250
QY 241 GAGGCGCTGAGCGCTGCTGCGGACAGACGACAGAGATCTGAATCAGGGCCCTGAGC 300
DB 251 GAGGCGCTGAGCGCTGCTGCGGACAGACGACAGAGATCTGAATCAGGGCCCTGAGC 310
QY 301 AAGCTGAGCTGCTCATCCGAGAAAGCGACAGAGCTTGCAGAGCTTACAGCGAGCACTGG 360

DB 311 AAGCTGAGCTGCTCATCCGAGAACGCGACAGCTTCCGAGACCTTACAGCGAGCATG 370
QY 361 CAGCAGCTGACAGAGGCTCAACAGACCAAGCCAGGACATTTGAGAGCTGAAGAC 420
DB 371 CAGCAGCTGACAGAGGCTCAACAGACCAAGCCAGGACATTTGAGAGCTGAAGAC 430
QY 421 CAGTACCGAGCTTGGCAGCGGACAGTGGCCCAAGCCAGGACATTTGAGAGCTT 480
DB 431 CAGTACCGAGCTTGGCAGCGGACAGTGGCCCAAGCCAGGACATTTGAGAGCTT 490
QY 481 AAGACAGAGCCGTTGACAGGACCAAGACCAAGTATGTCGAGCTGTGAAAGCTTT 540
DB 491 AAGACAGAGCCGTTGACAGGACCAAGACCAAGTATGTCGAGCTGTGAAAGCTTT 550
QY 541 GCTCACCAACCGCTATGCTGGGCGGCGGAGGCTGACAGTACCAACAGCACAC 600
DB 551 GCTCACCAACCGCTATGCTGGGCGGCGGAGCTGACAGTACCAACAGCACAC 610
QY 601 CACAGCTCTGCTGCGCGGCTGCTGCGGCTGCTGACAGACCTGACAGAGATGAGT 660
DB 611 CACAGCTCTGCTGCGCGGCTGCTGCGGCTGCTGACAGACCTGACAGAGATGAGT 670
QY 661 TGCATCTGAGAGAGATCTCTGACAGATATCTGAGATTTAGACAGCTTGTGAGATGAG 720
DB 671 TGCATCTGAGAGAGATCTCTGACAGATATCTGAGATTTAGACAGCTTGTGAGATGAG 730
QY 721 GTGGTGGCAATTCACCGGAGATGCTGAGAGCTGCTGCGGCTGACAGCTGAGCTGAG 780
DB 731 GTGGTGGCAATTCACCGGAGATGCTGAGAGCTGCTGCGGCTGACAGCTGAGCTGAG 790
QY 781 TACCAAGGCTTCTGCGACAGATATGAGGTCGACAGCTGACGCTGCTGCTGACGTTT 840
DB 791 TACCAAGGCTTCTGCGACAGATATGAGGTCGACAGCTGACGCTGCTGCTGACGTTT 850
QY 841 GATGAGTCACTGCTTGAAGAGGCTGAACGCTGAGAGCTTGGGAGCTTCAAGCTGAACGAG 900
DB 851 GATGAGTCACTGCTTGAAGAGGCTGAACGCTGAGAGCTTGGGAGCTTCAAGCTGAACGAG 910
QY 901 CTGAGCTGAGAGAGGCTGACAGACAGCTGACCTGAGTACAGATAGAGTGGCTGGCC 960
DB 911 CTGAGCTGAGAGAGGCTGACAGACAGCTGACCTGAGTACAGATAGAGTGGCTGGCC 970
QY 961 ACCGAGATGATGTTTACGCGGAGAGAGATGTTTACGAGCTGACAGAGAGCTCCGGAAT 1020
DB 971 ACCGAGATGATGTTTACGCGGAGAGAGATGTTTACGAGCTGACAGAGAGCTCCGGAAT 1030
QY 1021 GAAAGAGAAACACCCACCCCGGAGGCTGACAGCTGCTGAGGCAAGAGGCAAGTGTCTG 1080
DB 1031 GAAAGAGAAACACCCACCCCGGAGGCTGACAGCTGCTGAGGCAAGAGGCAAGTGTCTG 1090
QY 1081 CAAGAGACCTGACAGGAGGCTGACAGATGAGGCTGAGGCAAGAGGCAAGTGTCTG 1140
DB 1091 CAAGAGACCTGACAGGAGGCTGACAGATGAGGCTGAGGCAAGAGGCAAGTGTCTG 1150
QY 1141 CAGGAGTTCCTGAGACCAAGCTGAGACCTTGGGCGCCGCGAGGCTGTGTCTG 1200
DB 1151 CAGGAGTTCCTGAGACCAAGCTGAGACCTTGGGCGCCGCGAGGCTGTGTCTG 1210
QY 1201 CTCCTGAGAGATGACCGCACTTCCAGCTGCTCTGAGACAGAGAGGAGGAGGAG 1260
DB 1211 CTCCTGAGAGATGACCGCACTTCCAGCTGCTCTGAGACAGAGAGGAGGAGGAGGAG 1270
QY 1261 ACACCCAGCTGAGAGATCTTAAAGAGCAATCTGAGAACTTTCGCGCCCAAGTTCTG 1320
DB 1271 ACACCCAGCTGAGAGATCTTAAAGAGCAATCTGAGAACTTTCGCGCCCAAGTTCTG 1330
QY 1321 ----- 1320
DB 1331 CTCCTTCAACCGCTGAGCTCATTCGAGAGGTCAGAGAGCCCTGATGAGAGCTGTG 1390
QY 1321 ----- 1320
DB 1391 TACCAAGGAGCAATCCGAGGAGAGGAGAGTGGCTGAGCTGTGTGACCTTGGGAGCTTC 1450


```

1321 ----- 1320
1451 CTGGTGGGAGAGCCAGGAGAGAGTACGTGCTGGTGTGTGGATGTCTG 1510
1321 -----AACCTGACCGAAGGAGGAGGCTTT 1350
1511 CCCCAGCACTTATCATTCAGTCTTGGATTACCTGTACCGACTGGAGGGAGGCTTT 1570
1351 CCTAGCAATTCCTTGTCTATGACCACTATGAGACCCAGACGCCCTCAACGAAG 1410
1571 CCTAGCAATTCCTTGTCTATGACCACTATGAGACCCAGACGCCCTCAACGAAG 1630
1411 AGTGTGTGTCTCTGACAGGGCTGTGCCAAGACAACTGGGTGTGAACCATGAGAC 1470
1631 AGTGTGTGTCTCTGACAGGGCTGTGCCAAGACAACTGGGTGTGAACCATGAGAC 1690
1471 CTGGTGTGTGTGAGAGATTGGAACGGGGGAACTTTGGCGAAGTGTTCAGCGGACGCTG 1530
1691 CTGGTGTGTGTGAGAGATTGGAACGGGGGAACTTTGGCGAAGTGTTCAGCGGACGCTG 1750
1531 CGAGCCGACACACCCTGTGGCGGTGAAGTCTTGTGAGAGACGCTCCACCTGACCTC 1590
1751 CGAGCCGACACACCCTGTGGCGGTGAAGTCTTGTGAGAGACGCTCCACCTGACCTC 1810
1591 AAGGCCAAGTTTCTACAGAGACCGAGATCTTGAAGCACTACGCCAACCCCAATCGTG 1650
1811 AAGGCCAAGTTTCTACAGAGACCGAGATCTTGAAGCACTACGCCAACCCCAATCGTG 1870
1651 CCTCTCATTTGTGTCTGACAGCCAGAACGACCATCTATCATGCTATGAGAGCTTGTGAG 1710
1871 CCTCTCATTTGTGTGTCTGACAGCCAGAACGACCATCTATCATGCTATGAGAGCTTGTGAG 1930
1711 GGGGGGCACTTCTTGAACCTTCTCGACGAGAGGGGGCCGCTGCGGTGAAGACTGTG 1770
1931 GGGGGGCACTTCTTGAACCTTCTCGACGAGAGGGGGCCGCTGCGGTGAAGACTGTG 1990
1771 CTGCAAGATGTGGGGGATGACACTGTGTGCACTGAGATACCTGAGAGCAAGTCTGATC 1830
1991 CTGCAAGATGTGGGGGATGACACTGTGTGCACTGAGATACCTGAGAGCAAGTCTGATC 2050
1831 CACCGGAGCCTGAGCTGCTGGAATCTGCTGTGACAGAGAAATGCTCTGAAGATCACT 1890
2051 CACCGGAGCCTGAGCTGCTGGAATCTGCTGTGACAGAGAAATGCTCTGAAGATCACT 2110
1891 GACTTTGGGATGTCCCGAGAGAGAGCCGATGGGGTCTATGACAGCCTCAGGGGGCCTCAGA 1950
2111 GACTTTGGGATGTCCCGAGAGAGAGCCGATGGGGTCTATGACAGCCTCAGGGGGCCTCAGA 2170
1951 CAAATGCCCTGTGAGTGTGACCGGACCTGAGGCGCTTAACTAGCGCGCTACTCTCCGAA 2010
2171 CAAATGCCCTGTGAGTGTGACCGGACCTGAGGCGCTTAACTAGCGCGCTACTCTCCGAA 2230
2011 AAGCAGGTGTGAGAGCTTGGGATCTTGGCTGAGGAGACCTTCAAGCCTGAGGGGCTCCGCC 2070
2231 AAGCAGGTGTGAGAGCTTGGGATCTTGGCTGAGGAGACCTTCAAGCCTGAGGGGCTCCGCC 2290
2071 TATCCCAACTCAGCAATCAGACACACGAGATTGTGAGAGAGGGGGCGCTGCGCC 2130
2291 TATCCCAACTCAGCAATCAGACACACGAGATTGTGAGAGAGGGGGCGCTGCGCC 2350
2131 TGCCCAAGAGCTGTGTCTGATGCTCGGTTCAGGCTCATGAGAGAGTGTGGGCTTATGAG 2190
2351 TGCCCAAGAGCTGTGTCTGATGCTCGGTTCAGGCTCATGAGAGAGTGTGGGCTTATGAG 2410
2191 CCTGGGACAGGGGCCAGCTTCAAGACCATCTACAGAGAGCTGACAGACATCCGAAAGCGG 2250
2411 CCTGGGACAGGGGCCAGCTTCAAGACCATCTACAGAGAGCTGACAGACATCCGAAAGCGG 2470
2251 CATCGG 2256
2471 CATCGG 2476

```

RESULT 4

BC035357

2788 bp

mRNA linear PRI 29-JUN-2004

LOCUS

BC035357
 Homo sapiens feline sarcoma oncogene, mRNA (CDNA clone MGC:34142
 IMAGE:5170548), complete cds.

DEFINITION

BC035357.1 GI:23271524

ACCESSION

BC035357.1

VERSION

BC035357.1

KEYWORDS

MGC.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 2788)

AUTHORS

Strausberg, R. L., Pelting, E. A., Grouse, L. H., Berge, J. G., Klausner, R. D., Collins, F. S., Wagner, L., Shennan, C. M., Schuler, G. D., Altschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bhat, N. K., Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L., Stapleton, M., Soares, M. B., Bonaldo, M. P., Casavant, T. L., Scheetz, T. E., Brownstein, M. J., Ueda, T. B., Toshikuni, S., Carninci, P., Prange, C., Raha, S. S., Loquellano, N. A., Peters, G. J., Abramson, R. D., Mullan, S. J., Bosak, S. A., McEwan, P. J., McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S., Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulik, S. W., Villalón, D. K., Muzny, D. M., Sodergren, E. U., Lu, X., Gibbs, R. A., Fahey, J., Helton, E., Kettelman, M., Madan, A., C., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bouffard, G. G., Blakeley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M., Butterfield, Y. S., Krzywicki, M. I., Skalek, U., Small, D. E., Schermer, A., Schein, J. E., Jones, S. J., and Marra, M. A.

TITLE

Human and mouse cDNA sequences

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED

12477932

REFERENCE

2 (bases 1 to 2788)

AUTHORS

Strausberg, R.

TITLE

Direct Submission

JOURNAL

Submitted (31-JUN-2002)

REMARK

Contact: MGC help desk

COMMENT

Email: cgapbs-r@mail.nih.gov

REMARK

Tissue Procurement: Life Technologies, Inc.

REMARK

CDNA Library Preparation: Life Technologies, Inc.

REMARK

CDNA Library Arrayed by: Baylor College of Medicine Human Genome Sequencing Center

REMARK

Center code: BCM-HGSC

REMARK

Web site: http://www.hgsc.bcm.tmc.edu/cdna/

REMARK

Contact: amg@bcm.tmc.edu

REMARK

Gunaratne, P. H., Garcia, A. M., Lu, X., Hulik, S. W., Louieged, H., Kowis, G. R., Sneed, A. J., Martin, R. G., Muzny, D. M., Nanavati, A. N., Gibbs, R. A.

FEATURES

source

FEATURES

1..2788

FEATURES

/organism="Homo sapiens"

FEATURES

/mol_type="mRNA"

FEATURES

/db_xref="taxon:9606"

FEATURES

/clone="MGC:34142 IMAGE:5170548"

FEATURES

/tissue_type="Brain, adult medulla"

FEATURES

/clone_lib="Nih MGC_119"

FEATURES

/lab_host="DH10B"

FEATURES

/note="Vector: pCMV-SPORT6"

gene
1. .2788
/gene="FES"
/note="synonym: FES"
/db_xref="LocustID:2242"
/db_xref="MTM:190030"
73. .2541
/gene="FES"
/codon_start=1
/product="V-FES feline sarcoma viral/V-FPS fujinami avian
sarcoma viral oncogene homolog"
/protein_id="AAH35357.1"
/db_xref="GI:23271525"
/db_xref="LocustID:2242"
/db_xref="MTM:190030"
/translation="MGFSESLGSPGHGVLQMOEAEHLILSGMKMAORYKREY
AGLHHMSLDGSGGSRATSPSPISQSAEITISOTEGSLRLROAEPLNGSPSL
KLLIRKQRLRTYSQWQOOLQOELETKTHSDIEKLSQYRLARSDAKRYDAS
KOKDRKADKRYRSLMKLFAHNRVLGVRAQLHGHQHLFLGRYSADLHEE
MACILKEILQBYLEISLVQDEVAIHRMAAARIQEBAYOGFLRGYSAPVPP
CTPDSLEEGEPLEPGLQINELTVEVOHTLVTDLAVATAMVFRBDMVTL
QOELRREENTHPRERVOILGKROVALQGLQVALGQVALAQOELLQRLHEIG
RBPRLVLLQDHRHSTSSSEOREGAPPTLEIKSHISGIPRPFSLPPLQILPE
VQRLHEQLMTHAIPRAEVALVHSGPLVRESQKQBYLSVIMDLPHFTIIS
LNLVRLIEGEPSPILIDHLSTQPLTKSGVVLHRAVPDKVNLHEDLVGEQ
IGKNGEVEFSGRLRADNTLVAVKSCRETLPDLKAKFLQEBATILKQYHPNIVRLIG
VCTQKQPIYVIMELVQGDPLFLRTREGARLRYKTLQWVDAALAMEYLESKCIHR
DLAARNCIVTEKRVKISDFGMSREADGVAAAGLROVPVMTAPBALNTRYGYSSE
SPVMSGILLMTERSLGASPYPLSMQOTREFPEKGRPLPCELCDAVFRIMEQMA
YEPGQRPSPSTIYQELQSIKRRH"

ORIGIN

Query Match 90.2%; Score 2036; DB 9; Length 2788;

Best Local Similarity 91.5%; Pred. No. 0;

Matches 2256; Conservative 0; Mismatches 0; Indels 210; Gaps 1;

QY 1 ATGGGCTTCTCTTCTGAGCTGTGACGCCCCCAGGGCCACGGGGTCTCTGACGAAATGCA 60
DB 73 ATGGGCTTCTCTTCTGAGCTGTGACGCCCCCAGGGCCACGGGGTCTCTGACGAAATGCA 132
QY 61 GAGGCCGAGCTTGTCTACTGAGGGGCAATGAAAGTGAATGGCCAGCGGGTCAAGAGT 120
DB 133 GAGGCCGAGCTTGTCTACTGAGGGGCAATGAAAGTGAATGGCCAGCGGGTCAAGAGT 192
QY 121 GACAGGAGATGATCAGAGATGCTTCAACATGTCTCTGACAGAGATGGGGGCCAGAGC 180
DB 193 GACAGGAGATGATCAGAGATGCTTCAACATGTCTCTGACAGAGATGGGGGCCAGAGC 252
QY 181 CGGGCCATCAGCCCTGACAGCCCATCACTGCTGCTGGCTGATGATCACCAGCCAACT 240
DB 253 CGGGCCATCAGCCCTGACAGCCCATCACTGCTGCTGGCTGATGATCACCAGCCAACT 312
QY 241 GAGGGCTGAGCGGCTTGTCTGCGGCGAGCAAGAGATGTAATCAAGGCCCCCTGAGC 300
DB 313 GAGGGCTGAGCGGCTTGTCTGCGGCGAGCAAGAGATGTAATCAAGGCCCCCTGAGC 372
QY 301 AAGCTGAGCTGCTCATCCGGGAAACGGCAGACACTTGCAGAGCTTAAGAGCTGAG 360
DB 373 AAGCTGAGCTGCTCATCCGGGAAACGGCAGACACTTGCAGAGCTTAAGAGCTGAG 432
QY 361 CAGCAGCTGAGAGAGAGCTACCAAGACCCACAGCAGAGCAATTGAAAGCTGAAAGC 420
DB 433 CAGCAGCTGAGAGAGAGCTACCAAGACCCACAGCAGAGCAATTGAAAGCTGAAAGC 492
QY 421 CAGTACGAGACTCTGAGAGGGGACAGTGGCCAGGCAAGGCAAGTACCGAGAGGCGAGC 480
DB 493 CAGTACGAGACTCTGAGAGGGGACAGTGGCCAGGCAAGGCAAGTACCGAGAGGCGAGC 552
QY 481 AAAGACAAGAGCCGTGACAAAGCCAAAGACAAATATGTGCGACGCTGTGGAAGCTTTT 540
DB 553 AAAGACAAGAGCCGTGACAAAGCCAAAGACAAATATGTGCGACGCTGTGGAAGCTTTT 612
QY 541 GCTTACCAACACCGCTATGTGTGGGCTGCGGGCTGCGAGCTTACCAACACGACAC 600

DB 613 GCTCACCAACACCGCTATGTGTGGGCTGCGGGCTGCGACCTACCAACACGACAC 672
QY 601 CACAGAGCTCTGCTGCGCGGCTGCTGCGGCTCACTGACAGAGACTTGCAGAGAGT 660
DB 673 CACAGAGCTCTGCTGCGCGGCTGCTGCGGCTCACTGACAGAGACTTGCAGAGAGT 732
QY 661 TGCATCTGAAAGAGATCTGACAGAAATACCTGAGATTAGAGCTGTGTGACAGATGAG 720
DB 733 TGCATCTGAAAGAGATCTGACAGAAATACCTGAGATTAGAGCTGTGTGACAGATGAG 792
QY 721 GTGTGTGCGCATTCACCGGAGATGTGCTGACGCTGCCCGCATCCAGCTTGAAGCTGAG 780
DB 793 GTGTGTGCGCATTCACCGGAGATGTGCTGACGCTGCCCGCATCCAGCTTGAAGCTGAG 852
QY 781 TACCAAGGCTTCTGCGACAGATGAGGTTCGCACTGACGTCCACCTGTGTCACTTC 840
DB 853 TACCAAGGCTTCTGCGACAGATGAGGTTCGCACTGACGTCCACCTGTGTCACTTC 912
QY 841 GATGAGTCACTGCTTGAAGAGGCTGAACCGCTGAGACCTGGGAGCTCCAGCTGAACGAG 900
DB 913 GATGAGTCACTGCTTGAAGAGGCTGAACCGCTGAGACCTGGGAGCTCCAGCTGAACGAG 972
QY 901 CTGACTGTGAGAGAGCTGTGACACACGCTGACCTCAAGTACAGATGAGCTGTGTGCGC 960
DB 973 CTGACTGTGAGAGAGCTGTGACACACGCTGACCTCAAGTACAGATGAGCTGTGTGCGC 1032
QY 961 ACCGAGATGATGTTTACGCGGCGAGAGATGTTTACGCACTGTCACAGAGAGCTCCGGAAT 1020
DB 1033 ACCGAGATGATGTTTACGCGGCGAGAGATGTTTACGCACTGTCACAGAGAGCTCCGGAAT 1092
QY 1021 GAAAGAGAGAACACCCACCCCGGAGAGCGGCTGACGCTGTGGGCAAGGCAAGTGTCTG 1080
DB 1093 GAAAGAGAGAACACCCACCCCGGAGAGCGGCTGACGCTGTGGGCAAGGCAAGTGTCTG 1152
QY 1081 CAAGAGACCTGACAGAGGCTGTGAGAGTGAAGGCTGTGACGAGGCAAGGCTGACGAGCCAG 1140
DB 1153 CAAGAGACCTGACAGAGGCTGTGAGAGTGAAGGCTGTGACGAGGCAAGGCTGACGAGCCAG 1212
QY 1141 CAGAGATTGCTGACAGACCAAGCTGTGAGCACTGTGGGCCCGGAGAGCCCGCTGTGTCTG 1200
DB 1213 CAGAGATTGCTGACAGACCAAGCTGTGAGCACTGTGGGCCCGGAGAGCCCGCTGTGTCTG 1272
QY 1201 CTCTTGCAGAGATGACCGGCACTTCCACTGTCTCTGAGACAGAGGCGAGAGGCGGAGAGG 1260
DB 1273 CTCTTGCAGAGATGACCGGCACTTCCACTGTCTCTGAGACAGAGGCGAGAGGCGGAGAGG 1332
QY 1261 ACAACCAAGCTGAGATCTCTTAAAGGCCACATCTCAAGAACTCTTCCGCCCAAGTCTG 1320
DB 1333 ACAACCAAGCTGAGATCTCTTAAAGGCCACATCTCAAGAACTCTTCCGCCCAAGTCTG 1392
QY 1321 ----- 1320
DB 1393 CTCCCTCAACCGCTGAGCTCATTCGAGAGGTGCAGAAACCCCTGTGATGAGACGTGTG 1452
QY 1321 ----- 1320
DB 1453 TACACAGGGGCAATCCAGAGGAGAGGTGTGAGCTGTGATGATCTGTGGGCACTTC 1512
QY 1321 ----- 1320
DB 1513 CTGTGTGCGGAGAGCCAGGAGCAAGAGATACGTGTCTGTGTGTGTGATGTCTG 1572
QY 1321 -----AACTGTACCGACTGTGAAGGGAGGCTTT 1350
DB 1573 CCGCGGCACTTCATCATTCAGTCCCTTGATTAACCTGTACGACTGAGAGGGAGGCTTT 1632
QY 1351 CCTAGCATTCCTTTGTCTATTCAGCACTTGTAGCACTTGTAGCACTTGTAGCACTTGTAGCA 1410
DB 1633 CCTAGCATTCCTTTGTCTATTCAGCACTTGTAGCACTTGTAGCACTTGTAGCACTTGTAGCA 1692
QY 1411 AGTGTGTGTGTCTGTGACAGGCTGTGCGCAAGAGCAATGTGTGTGTGAACCATGAGGAC 1470
DB 1693 AGTGTGTGTGTCTGTGACAGGCTGTGCGCAAGAGCAATGTGTGTGTGAACCATGAGGAC 1752

[illegible]

JOURNAL	Identification of the 5' end of the gene
MDLINE	Oncogene 5 (3), 267-275 (1990)
PUBMED	90191711
FEATURES	2179816
source	Location/Qualifiers
gene	1..2771
CDS	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	1..2771
	/gene="c-fes"
	76..2544
	/gene="c-fes"
	/cdon_start=1
	/protein_id="CAA6438.1"
	/db_xref="GI:29891"
	/db_xref="GOA:P07332"
	/translation="MGFSESLCPQHGVLQOQOEALRLLEGKMKMAORVKSREK
	AGLHHMSLSDSGSOSALIPDPSISQSMELTSCGSLRLROHAPEDLNSGPLSKL
	SLLRQDLKRTSIEWQQLQDELITKHSODIEKLKSQYRALLRDSQAKRKIQBAS
	KIDROKADKRVKSLMKLPAHNRYVLGVAAQLNHQHHQLLPLGLRLQDLHEE
	MACLKELIQEYLEISLQDEVVAIHREMAAARLQPEAYEGFPLWFGSADVPK
	CTFDESLLEEGPELSEGLQNLNLFVESYQHTLTSVDELAVAIEMFRQESADP
	QOELRNEENTHPRERYQLLQKQVLOALQVALCSQAKLQAOELLOTLEHGL
	PGEPPPLLLQDDRHSITSSSEQEREGRTPLLELKSHTLSQIRPKSLPPLQIPE
	VQKRLHQMLHYGAIIPRAEVAELVHSGDPLVRSQSKQIEVLSLMDGLRHLFIQ
	LDMVRLREGGFPSPILLIDHLSTQDPLTRKSSGVNLAAPKQKWLNHDLVLEGEC
	IGRNFGEVFSGRRLADNTLVAVKSCHETPLPDKAFLQEARILKOYSHNIVALLIG
	VCQKQPIYIVMELVOGDFLPLRTGARGRLVYTLTLOWGDAAGNEYLESKCIH
	DLAARNCLYTEKVTIKTSDRCMSREADGVYLAASGSRQVPMKTPAPALNYGRSS
	SDWSRGITLMEITSLGASPPNLSNQOTREVEKGRLLPCELPDAVFLMEQCM
	YEPQRPSTITLYQELQSIKKRHH"
	2738..2743
	/gene="c-fes"
polya_signal	
ORIGIN	
Query Match	90.0%; Score 2029.6; DB 9; Length 2771;
Best Local Similarity	91.3%; Pred. No. 0;
Matches 2252; Conservative	0; Mismatches 4; Indels 210; Gaps 1.
QY	1 ATGGGCTTCCTCTGAGCTGTGAGGCCCCAGGGGCCACGGGGTCTTGACCAATGACG 60
Db	76 ATGGGCTTCCTCTCGAGCTGTGAGGCCCCAGGGGCCACGGGGTCTTGACCAATGACG 135
QY	61 GAGGCCGAGCTTCGCTACTGTGAGGGGATGAGAAAGTGATGGCCAGCGGGTCAAGAGT 120
Db	136 GAGGCCGAGCTTCCTACTGTGAGGGGATGAGAAAGTGATGGCCAGCGGGTCAAGAGT 195
QY	121 GACAGGGAGTATGAGAGACTGCTTCACCACTGTCCTTGACAGACAGTGGGGGCCAGAGC 180
Db	196 GACAGGGAGTATGAGAGACTGCTTCACCACTGTCCTTGACAGACAGTGGGGGCCAGAGC 255
QY	181 CGGGCCATCAGCCCTTGACAGGCCCATCATAGTACTTCGGGTGAGATACACAGCCAACT 240
Db	256 CGGGCCATCAGCCCTTGACAGGCCCATCATAGTACTTCGGGTGAGATACACAGCCAACT 315
QY	241 GAGGGCCTGAGCCGCTTGCTGCGGACAGACGCAAGAGATCTGAATCTCAGGGCCCTTGAGC 300
Db	316 GAGGGCCTGAGCCGCTTGCTGCGGACAGACGCAAGAGATCTGAATCTCAGGGCCCTTGAGC 375
QY	301 AAGCTGAGCTGTCTCATCCGGGAAACGGACGAGCTTGCAAGACTTACAGGAGCAGTGG 360
Db	376 AAGCTGAGCTGTCTCATCCGGGAAACGGAGAGCTTGCAAGACTTACAGGAGCAGTGG 435
QY	361 CAGCAGCTGCAGCAGAGCTCAACCAAGACCTCAAGCCAGCAATTGAGAACTGAAGGC 420
Db	436 CAGCAGCTGCAGCAGAGCTCAACCAAGACCTCAAGCCAGCAATTGAGAACTGAAGGC 495
QY	421 CAGTACCGAGCTCGGACACGGGACAGTGGCCCAAGCGCAAGTACAGAGAGCCAGC 480
Db	496 CAGTACCGAGCTCGGACACGGGACAGTGGCCCAAGCGCAAGTACAGAGAGCCAGC 555

QY 481 AAAGACAGGACCGGTGACAGGACCAAGATATGTGCGACGCTGTGGAAGCTCTT 540
DB 556 AAAGACAGGACCGGTGACAGGACCAAGATATGTGCGACGCTGTGGAAGCTCTT 615
QY 541 GCTTACACCAACCGCTATGTGTGCGGTGCGGCTGCGGACTTACACCAACGACAC 600
DB 616 GCTTACACCAACCGCTATGTGTGCGGTGCGGCTGCGGACTTACACCAACGACAC 675
QY 601 CACGAGCTCTGCGCGCGGCTGCGGCTGCGGCTGCGGACTTACACCAACGACAC 660
DB 676 CACGAGCTCTGCGCGCGGCTGCGGCTGCGGCTGCGGACTTACACCAACGACAC 735
QY 661 TGCATCTGGAAGAGATCTGCAAGATATCTGCAAGATATGCAAGCTGTGCAAGAT 720
DB 736 TGCATCTGGAAGAGATCTGCAAGATATCTGCAAGATATGCAAGCTGTGCAAGAT 795
QY 721 GTGTGTGCGCTTACACGCGGAGATGTGCGGCTGCGGCTGCGGACTTACACCAAC 780
DB 796 GTGTGTGCGCTTACACGCGGAGATGTGCGGCTGCGGCTGCGGACTTACACCAAC 855
QY 781 TACCAAGGCTTCTGCGGACAGATATGAGTCTGCGGCTGCGGACTTACACCAAC 840
DB 856 TACCAAGGCTTCTGCGGACAGATATGAGTCTGCGGCTGCGGACTTACACCAAC 915
QY 841 GATGAGTCACTGCTTGAAGAGAGTGAACGCTGAGGCTGCGGAGCTTACAGTGAAC 900
DB 916 GATGAGTCACTGCTTGAAGAGAGTGAACGCTGAGGCTTACAGTGAACGAGTGAAC 975
QY 901 CTGATCTGTGAGAGCGTGTGCAACACGCTGACCTCACTGATGACATGAGTGTGAGCC 960
DB 976 CTGATCTGTGAGAGCGTGTGCAACACGCTGACCTCACTGATGACATGAGTGTGAGCC 1035
QY 961 ACCGAGTGTGTTGACGCGGAGAGATGATTACGAGCTGCAACGAGGCTCCGGAAT 1020
DB 1036 ACCGAGTGTGTTGACGCGGAGAGATGATTACGAGCTGCAACGAGGCTCCGGAAT 1095
QY 1021 GAAAGAGAGAACCCACCCCGGAGAGCGGAGTCAAGCTGTGAGGCAAGAGTGTG 1080
DB 1096 GAAAGAGAGAACCCACCCCGGAGAGCGGAGTCAAGCTGTGAGGCAAGAGTGTG 1155
QY 1081 CAAAGAGACATGCAAGGAGGCTGTGCAAGTGTGCAAGGCTGTGCAAGGCTGTG 1140
DB 1156 CAAAGAGACATGCAAGGAGGCTGTGCAAGTGTGCAAGGCTGTGCAAGGCTGTG 1215
QY 1141 CAGAGTGTGTTGACGCGGAGAGATGATTACGAGCTGCAACGAGGCTCCGGAAT 1200
DB 1216 CAGAGTGTGTTGACGCGGAGAGATGATTACGAGCTGCAACGAGGCTCCGGAAT 1275
QY 1201 CTCTGTGAGATGACCGGCACTTCAAGTGTGCTGTGAGGAGAGAGGAGGAGG 1260
DB 1276 CTCTGTGAGATGACCGGCACTTCAAGTGTGCTGTGAGGAGAGAGGAGGAGG 1335
QY 1261 ACACCCACGCTGAGAGATCTTAAAGGACCAATCTCAAGGAATCTTCCGCCCCAAGTTCTG 1320
DB 1336 ACACCCACGCTGAGAGATCTTAAAGGACCAATCTCAAGGAATCTTCCGCCCCAAGTTCTG 1395
QY 1321 ----- 1320
DB 1396 CTCTGTGAGATGACCGGCACTTCAAGTGTGCTGTGAGGAGAGGAGGAGG 1455
QY 1321 ----- 1320
DB 1456 TACCAAGGCTTCTGCGGACAGATATGAGTCTGCTGATGACCTGTGAGGACTTC 1515
QY 1321 ----- 1320
DB 1516 CTGTGTGCGGAGAGCGGAGAGAGAGATAGTGTGCTGTGAGGAGTGTGCTG 1575
QY 1321 ----- 1320
DB 1576 CCGGCGGACCTTCACTACAGTCTTGTGATTAAGTGTGAGAGGAGGAGGCTTT 1635
QY 1351 CTAAGCATTCCTTGTCTATGACCACTTCTGAGCAACCGAGAGCCCTTACCAAGAAAG 1410

DB 1636 CTAAGCATTCCTTGTCTATGACCACTTCTGAGCAACCGAGAGCCCTTACCAAGAAAG 1695
QY 1411 AGTGTGTGTGCTGTCAGACGAGGCTGTGAGCCCAAGGAGCAAGTGGGTGTGAACATGAGAC 1470
DB 1696 AGTGTGTGTGCTGTCAGACGAGGCTGTGAGCCCAAGGAGCAAGTGGGTGTGAACATGAGAC 1755
QY 1471 CTGTGTGTGTGAGCAGATTTGAGAGGAGGAACTTTTGGGAAAGTGTTCAGCGAGCGCTG 1530
DB 1756 CTGTGTGTGTGAGCAGATTTGAGAGGAGGAACTTTTGGGAAAGTGTTCAGCGAGCGCTG 1815
QY 1531 CGAGCCGACACACCTGTGTGCGGTGAAGTCTTGTGAGAGACGCTCCCACTGACCTC 1590
DB 1816 CGAGCCGACACACCTGTGTGCGGTGAAGTCTTGTGAGAGACGCTCCCACTGACCTC 1875
QY 1591 AAGGCGCAAGTTCTACAGAGACGAGATCTGAGCAATACAGCCCAACATCGTG 1650
DB 1876 AAGGCGCAAGTTCTACAGAGACGAGATCTGAGCAATACAGCCCAACATCGTG 1935
QY 1651 CGTCTCATTTGTGTGTCAGACCGGAGAGAGCCCATCTACATGATGATGAGGCTTGTGAG 1710
DB 1936 CGTCTCATTTGTGTGTCAGACCGGAGAGAGCCCATCTACATGATGATGAGGCTTGTGAG 1995
QY 1711 GGGGCGGACCTTCTGACCTTCTCGGACGAGAGGAGGCGGCTGCGGAGTGAAGTCTG 1770
DB 1996 GGGGCGGACCTTCTGACCTTCTCGGACGAGAGGAGGCGGCTGCGGAGTGAAGTCTG 2055
QY 1771 CTGCAAGTGTGAGGAGATGCACTGTGCAAGGATGATCTGGAAGAGAGTGTGAGT 1830
DB 2056 CTGCAAGTGTGAGGAGATGCACTGTGCAAGGATGATCTGGAAGAGAGTGTGAGT 2115
QY 1831 CACCGGAGCCTGCTGCTGTGCAAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGT 1890
DB 2116 CACCGGAGCCTGCTGCTGTGCAAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGT 2175
QY 1891 GACTTGTGAGATGCTTGTGAGAGAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1950
DB 2176 GACTTGTGAGATGCTTGTGAGAGAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2235
QY 1951 CAAAGTCCCGTGAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2010
DB 2236 CAAAGTCCCGTGAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2295
QY 2011 AGGAGCTGTGAGGCTTGTGCAAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGT 2070
DB 2296 AGGAGCTGTGAGGCTTGTGCAAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGT 2355
QY 2071 TATCCCACTTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2130
DB 2356 TATCCCACTTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2415
QY 2131 TGCCCAAGCTGTGCTGATGCGGAGTCAAGGCTCATGAGAGAGTGTGAGGCTTGTGAG 2190
DB 2416 TGCCCAAGCTGTGCTGATGCGGAGTCAAGGCTCATGAGAGAGTGTGAGGCTTGTGAG 2475
QY 2191 CTTGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2250
DB 2476 CTTGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2535
QY 2251 CATCGG 2256
DB 2536 CATCGG 2541

RESULT 6
AX281847
LOCUS AX281847 2889 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 256 from Patent W00177389.
ACCESSION AX281847
VERSION AX281847.1 GI:16609098
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Source

1 Shiffman, D., Somogyi, R., Lawn, R., Sellhame, J.J., Porter, G.J., Mikkil, T. and Tai, J.
Genes expressed in foam cell differentiation
Patent: WO 017389-A 256 18-OCT-2001;
Incyte Genomics, Inc. (US)
Location/Qualifiers
1. 2889
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 997347.6"

ORIGIN

Query Match 89.7%; Score 2023.4; DB 6; Length 2889;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 2255; Conservative 0; Mismatches 1; Indels 211; Gaps 2;

QY 1 ATGGGCTTCTTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTTGCAGCAATGCAAG 60
DB 188 ATGGGCTTCTTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTTGCAGCAATGCAAG 247
QY 61 GAGGCGAGCTTGTCTACTGAGGCGATGAGAAAGTGGCCAGAGGAGTCAAGAT 120
DB 248 GAGGCGAGCTTGTCTACTGAGGCGATGAGAAAGTGGCCAGAGGAGTCAAGAT 307
QY 121 GAGGCGAGTATGAGAGTCTTCAACATGTCTTGCAGAGCAAGTGGGCGCAAGC 180
DB 308 GACAGGAGATGAGAGTCTTCAACATGTCTTGCAGAGCAAGTGGGCGCAAGC 367
QY 181 CGGCGCATCAGCCCTGACAGCCCATCATGCTCTGGGCTGAGATCACAGCCAACT 240
DB 368 CGGCGCATCAGCCCTGACAGCCCATCATGCTCTGGGCTGAGATCACAGCCAACT 427
QY 241 GA-GGGCTGAGCGGCTGTGCGGAGCAAGGAGATGTGAATCAAGGCGGCTGAG 299
DB 428 GAGGCGCTGAGCGGCTGTGCGGAGCAAGGAGATGTGAATCAAGGCGGCTGAG 487
QY 300 CAGCTGAGCTTGTCTATCGGAGCAAGGAGCTTGCAGAGCACTTACAGGAGAGTG 359
DB 488 CAGCTGAGCTTGTCTATCGGAGCAAGGAGCTTGCAGAGCACTTACAGGAGAGTG 547
QY 360 GCAAGCTGCTGAGAGAGTCTTCAACAGCCCAAGGAGATGAGAGTGAAGAG 419
DB 548 GCAAGCTGCTGAGAGAGTCTTCAACAGCCCAAGGAGATGAGAGTGAAGAG 607
QY 420 CAGCTGAGCTTGTGAGAGGAGAGTCCAGAGCAAGGAGATGAGAGGAGCAG 479
DB 608 CAGCTGAGCTTGTGAGAGGAGAGTCCAGAGCAAGGAGATGAGAGGAGCAG 667
QY 480 CAAAGCAAGAGAGCTGAGCAAGGAGCAAGAGTATGTGCGAGCTGTGAGAGCTTT 539
DB 668 CAAAGCAAGAGAGCTGAGCAAGGAGCAAGAGTATGTGCGAGCTGTGAGAGCTTT 727
QY 540 TGTCTACCAACACGCTATGTGCTGGGCGTGGCGGCTGCGAGCTTACCAACAGCA 599
DB 728 TGTCTACCAACACGCTATGTGCTGGGCGTGGCGGCTGCGAGCTTACCAACAGCA 787
QY 600 CCAAGGAGCTTGTGAGAGGAGAGTCTGAGAGAGTCAAGGAGATGAGAGTGGC 659
DB 788 CCAAGGAGCTTGTGAGAGGAGAGTCTGAGAGAGTCAAGGAGATGAGAGTGGC 847
QY 660 TTGCATCTGAGAGAGATCTTGAAGAGATCTGAGAGTGAAGGCTGTGAGAGTGA 719
DB 848 TTGCATCTGAGAGAGATCTTGAAGAGATCTGAGAGTGAAGGCTGTGAGAGTGA 907
QY 720 GGTGTGGGCTTCAACCGGAGAGTGGCTGAGCTGTGCGGCTGAGCTGAGCTGA 779
DB 908 GGTGTGGGCTTCAACCGGAGAGTGGCTGAGCTGTGCGGCTGAGCTGAGCTGA 967
QY 780 GTACCAAGGCTTCTGAGAGAGTATGAGTCCGAGCACTGAGCTCCACCTGTGTACGTT 839

DB 968 GTACCAAGGCTTCTGAGAGAGTATGAGTCCGAGCACTGAGCTCCACCTGTGTACGTT 1027
QY 840 CCAATGATCATCTGCTTGAAGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGAACGA 899
DB 1028 CCAATGATCATCTGCTTGAAGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGAACGA 1087
QY 900 GCTGAGTGTGAGAGGAGTGAACCAACCGCTGAGAGCTTGGGAGCTTCAAGTGAACGA 959
DB 1088 GCTGAGTGTGAGAGGAGTGAACCAACCGCTGAGAGCTTGGGAGCTTCAAGTGAACGA 1147
QY 960 CACCGAGATGTGTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGAACGA 1019
DB 1148 CACCGAGATGTGTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGAACGA 1207
QY 1020 TGAAGAGAGAACCAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1079
DB 1208 TGAAGAGAGAACCAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1267
QY 1080 GCAAGAGAGAGTGAAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1139
DB 1268 GCAAGAGAGAGTGAAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1327
QY 1140 GCAAGAGAGTGAAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1199
DB 1328 GCAAGAGAGTGAAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1387
QY 1200 GCTCTGAGAGTGAAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1259
DB 1388 GCTCTGAGAGTGAAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1447
QY 1260 GACACCAAGAGTGAAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1319
DB 1448 GACACCAAGAGTGAAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1507
QY 1320 G----- 1320
DB 1508 GCTCCTTCAACCGCTGAGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1567
QY 1321 ----- 1320
DB 1568 GTACCAAGGAGTGAAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1627
QY 1321 ----- 1320
DB 1628 CTTGTGCGGAGAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1687
QY 1321 ----- 1320
DB 1688 GCTCCTTCAACCGCTGAGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1747
QY 1350 TCTTATGATCTTCTTGTGATGAGCAACCTTATGAGCAACCTTATGAGCAACCTTATGAGCAACCTTATGAGCA 1409
DB 1748 TCTTATGATCTTCTTGTGATGAGCAACCTTATGAGCAACCTTATGAGCAACCTTATGAGCAACCTTATGAGCA 1807
QY 1410 GAGTGTGTGTGCTGAGCAAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1469
DB 1808 GAGTGTGTGTGCTGAGCAAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1867
QY 1470 CTTGTGAGTGAAGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1529
DB 1868 CTTGTGAGTGAAGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1927
QY 1530 GCGAGCGAGCAACCAACCTGAGAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1589
DB 1928 GCGAGCGAGCAACCAACCTGAGAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1987
QY 1590 CAAAGGCAAGTGTGAGAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1649
DB 1988 CAAAGGCAAGTGTGAGAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 2047
QY 1650 GGTCTCATTTGTGTGAGCAACCGAGGAGTGAACCGCTGAGAGGAGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGA 1709

Db		2648	GCGHCTCATTTGTCGTCCAGACCCGAGAAGCAGACCCTTACATGCATGAGAGCTTGTCGA	2107
Oy		1710	GGGGGGGCGACTTCTTGACCTTCCTCCGACGAGAGGGGGCCCCGCTGGGGGTGAAAGACTCT	1769
Db		2108	GGGGGGGCGACTTCTTGACCTTCCTCCGACGAGAGGGGGCCCCGCTGGGGGTGAAAGACTCT	2167
Oy		1770	GCTCAGATGGATGGAGGAGATGCAAGCTGCTGGGATGAGATCACTGGAGACAAGTCTGCAT	1829
Db		2168	GCTCAGATGGATGGAGGAGATGCAAGCTGCTGGGATGAGATCACTGGAGACAAGTCTGCAT	2227
Oy		1830	CCACCGGAGCTTGGCTGCTCGGAACTGCTGGATGACAGAGAAGATGTCCTGAAGATCAG	1889
Db		2228	CCACCGGAGCTTGGCTGCTCGGAACTGCTGGATGACAGAGAAGATGTCCTGAAGATCAG	2287
Oy		1890	TGACTTTTGGATGTGCCGAGAGAACCCGATGGGGTCTATTGACGCTTCAGGGGGCTTCAG	1949
Db		2288	TGACTTTTGGATGTGCCGAGAGAACCCGATGGGGTCTATTGACGCTTCAGGGGGCTTCAG	2347
Oy		1950	ACAAGTCCCCTGTAAGTGGACCGCACCTGAGAGCCCTTAACATAGGCGGCTACTCTCCGA	2009
Db		2348	ACAAGTCCCCTGTAAGTGGACCGCACCTGAGAGCCCTTAACATAGGCGGCTACTCTCCGA	2407
Oy		2010	AAGGACGCTGTGAGCTTTGGCATCTTGCTCTGGAGACCTTTCAGGCTGGGGGCTCCCC	2069
Db		2408	AAGGACGCTGTGAGCTTTGGCATCTTGCTCTGGAGACCTTTCAGGCTGGGGGCTCCCC	2467
Oy		2070	CTATCCCAACTCAGCAATTCAGACAGACACGGGAGATTGTGAGAAAGGGGGGCGCTTCGC	2129
Db		2468	CTATCCCAACTCAGCAATTCAGACAGACACGGGAGATTGTGAGAAAGGGGGGCGCTTCGC	2527
Oy		2130	CTGGCCAGAGGTGTGCTCGTGANGCCGTGTTCAGGCTCATGTGAGACAGTCTGGGCGCTAAGA	2189
Db		2528	CTGGCCAGAGGTGTGCTCGTGANGCCGTGTTCAGGCTCATGTGAGACAGTCTGGGCGCTAAGA	2587
Oy		2190	GCCTGGGAGAGGGGCCACAGCTTCAGACCATCTAACAGAGAGCTGACAGACATCCGAAACGG	2249
Db		2588	GCCTGGGAGAGGGGCCACAGCTTCAGACCATCTAACAGAGAGCTGACAGACATCCGAAACGG	2647
Oy		2250	GCATCGG 2256	
Db		2648	GCATCGG 2654	
RESULT 7				
MPEBSR				
LOCUS				
DEFINITION			2680 bp mRNA linear ROD 05-JUN-1995	
ACCESSION			Mouse c-fes proto-oncogene mRNA for c-fes protein.	
VERSION			X12616	
KEYWORDS			X12616.1 GI:50955 fes cellular oncogene; fes proto-oncogene; proto-oncogene; tyrosine kinase.	
SOURCE				
ORGANISM			Mus musculus (house mouse)	
			Mus musculus	
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE				
AUTHORS			1 (bases 1 to 2680)	
TITLE			Wilks,A.F. and Kurban,R.R.	
JOURNAL			Isolation and structural analysis of murine c-fes CDNA clones	
MEDLINE			Oncogene 3 (), 289-294 (1988)	
PUBMED			89083198	
FEATURES				
source			Location/Qualifiers	
			1..2680	
			/organism="Mus musculus"	
			/mol_type="mRNA"	
			/db_xref="taxon:10090"	
			/clone_lib="lambda gcl1."	
			60..2522	
			/note="unamed protein product; c-fes protein"	
			/codon_start=1	
			/protein_id="CAA31138.1"	
			/db_xref="GI:50955"	
			/db_xref="GDB:P16879"	
CDS				

Query Match	Best Local Similarity	68.3%: Score 1540.4; DB 10; Length 2680;	Matches 1954; Conservative 0; Mismatches 296; Indels 216; Gaps 2
Qy	1	ATGGGCTTCTCTTCTGAGCTGTGTGAGCCCGCCAGGGGCCACGGGGTCTTGACGCAAAATGAGC	60
Db	60	ATGGGCTTCTCTTCTGAGCTGTGTGAGCCCGCCAGGGGCCACGGGGGAGTGTGACAGATGAGC	119
Qy	61	GAGGCCAGCTTGGTCTTACTGAGAGGCATGAGAAATGAGATGGCCCAAGCGGGTCAAGAGT	120
Db	120	GAACCTAGCGTCCCTTATTTGAGAGGCAATGAGAAATGAGATGGCCCAAGAGGTCAAGAGT	179
Qy	121	GACAGGGAGATATGAGAGAGCTTCAACAACATGTCCCTGACAGGACAGTGGGGGCCAGAGC	180
Db	180	GACCGGAAATATGAGAGATTTGCTTCAACAATGTCTTTGACAGACAGCGGAGGCCAGAGC	239
Qy	181	CGGGCCATCAAGCCCTGTACAGGCCCATCATGATCTGTGGGCTGAGATCAACAGCCAACT	240
Db	240	TGAGAGAGCGGGCCCCGACAGCCCTGTACAGCCAGTCTGTGGGACAGATTAACAAGCCAGACC	299
Qy	241	GAGGGCTGAGCGGCTTGTCTGCGGACAGCAGCAAGAGATCTGAATCAAGGCCCTCTGAGC	300
Db	300	GAGAACTTGAAGCCCGGGTGTCTGCGGACAGATCAAAAGATCTGAATCTGGGGCCCTTGAAGC	359
Qy	301	AAGCTGAGCCGTCTCATCCGGGAAACGGGACAGGCTTGGCAAGCCTTACAGCAGCAGAGTGG	360
Db	360	AAACTGAGCCTGTGATCCGGGAGCGGACAGCCTTGAGAAAGACCTTACAAGAGCAGTGG	419
Qy	361	CAGCAGCTGACAGCAGAGCTCACCAAGACCACAGCCAGACATTTGAGAAAGCTGAAGAC	420
Db	420	CAGCAGCTGACAGCAGAGCTCACCAAGACCACAGCCAGCAGCATTTGAGAAAGCTGAAGACT	479
Qy	421	CAGTACCGAGCTGTGCAACGGGACAGTGTCCCAAGCCCAAGCCCAAGTACCAAGAGGCCAGC	480
Db	480	CAGTACCGGACCTCTGTGATCGAGATGACACCCAGGCGCAGGCCCAAGTACCAAGAACCCAGC	539
Qy	481	AAABAACAAGACCGGTGCAAGGCCAAGGACAAAGTATGTGGCAGCGCTGTGGAAGCTCTTT	540
Db	540	AAAACAAGAATTCAGACCAAGGCCAAGGACAAAGTATGTCCGAGGCTGTGGAAGCTCTTT	599
Qy	541	GCTCACAACAACCGCTATGTGTCTGGGCGCTGCGGGCTGTGCGAGTTACACCAAGCAGCAGC	600
Db	600	GCCCAACAACAACCGCTATGTGTCTGGGCGCTGCGGGCGCGGACAGCTGTGACCAACAACAAC	659
Qy	601	CACCAAGCTCTGTCTGCCCGGCGCTGTGCGGTGATCTGACAGAACCTTGACAGAGAGATGGCT	660
Db	660	CACCGGTTCAATGTCTGTGCGCTGTGCGGTGATCTGACAGAACCTTGACAGAGAGATGGCG	719
Qy	661	TGCAATCTGAAGAGATCTTGTGAGAAATCTTGAAGATTACAGCCTGTGTGAGAGATGAG	720
Db	720	GGAATTTGGAAGACATCTGTGAGAAATCTTGAAGATTACAGCTGTGTGAGAGAGAT	779
Qy	721	GTGCTGGCAATTCACCGGAGATGGCTGACCTGTGCGCGCATTCACGCTTGAGGCTGAG	780
Db	780	GTGGCAATTCATTCACCGGAGCTGCTGACGCTGTGCTGTGAGATTCACGCTTGAGGCTTGAAG	839

QY	781	TACCAAGGCTTCCGCGACAGTAATAGGGTCCGCACTGACGCTCCACCCGTGTCACTGTC	840
Db	840	TACTTAGGCTTCCGCGACATATATGATATCCACCCTGATATGCGCACTGTGTCACTTTT	899
QY	841	GATGATCACTGCTTTGAGAGAGGTGAACCGCTGGAAGCTTGAGAGCTCACTGAACGAG	900
Db	900	GATGATGCTGCTTTGTAAGAAGGGGAACAGCTGAGGCCAGGGGAGCTCACTGAATGAG	959
QY	901	CTGACTGTGAGAGACGTGACACACGCTTACCTCAGTGAACATATGACTGTGTGTGCC	960
Db	960	TTGACTCTTGAGACGTGTCAACACGCTGACTTCTGTGAACATGAATCTGCTGTGCC	1019
QY	961	ACCGAGATGTGTTTCAAGCGCGCAGAGAGATGTTATCGCAGCTGCAACAGAGACTCCGAA	1020
Db	1020	ACCAAGAGGTGTGTGAGCCGCGCAGAGATGTGTCAGTCACTGTGACGCTGTGAACTCCAAAG	1079
QY	1021	GAAAGAGAAACACCCACCCCCGGGAGCGGTGTGACTGTCTGTGGCAGAGGCMAAGTGTG	1080
Db	1080	GAGAAACGAACACCCACCCCCGGGAGCGGTGTGACTGTCTGTAGCCMAAGGCAAGTGTG	1139
QY	1081	CAAGAAGCACTGCAGGGGCTCAGGTATCGCTGTGCAGCCAGGCGCAAGCTCAGAGCCAG	1140
Db	1140	CAAGAAGCAATACAGAGGTTTCAATATGACATGTGTGACGCAAGACMAAGCTTCAGGCCAG	1199
QY	1141	CAGAGTTTGTGCAGACCAACAGCTGAGACCTTGAGCCCGCGCGAGCCCGCTGTGTG	1200
Db	1200	CAGAACTTCTGCAGAGACGATGTGAGCAGCTCGGCAAGCGGCGAGCCCGCAGCGTGTG	1259
QY	1201	CTCCTGACGATGACCGGCACTCCACGTGTGTCTTCGAGACAGGAGCGAGAGGGGGAAAG	1260
Db	1260	CTCTTTCAGATGACCGCATTTCCACTCTTCAC- - -GAGCGAGAGAGGGGAAAG	1313
QY	1261	ACACCCAAGCTGAGATCTCTTAAGAGCCACATCTCAGGAATCTTCGCGCCCAAGTTCTC	1319
Db	1314	ACACCCAAGCTGAGATCTCTTAAGAGCCACATCTTCGGAATCTTCGCGCCCAAGTTCTC	1373
QY	1320	- - - - -	1319
Db	1374	ATCCCTCCACCACTGACAGCTGTGCGGAGGTGACAAAGCCTCTTATAGCAGCTGTG	1433
QY	1320	- - - - -	1319
Db	1434	TACCAAGGAGTATCCCTGTGGCAGAGGTACTGAGCTGTAAAGCACTGGGAGCTTC	1493
QY	1320	- - - - -	1319
Db	1494	CTGTTGCGAGAGACGAGGCAAGCAGAGATGTAATGTCGTGTGATGTGGATATGCCAC	1553
QY	1320	- - - - -GAACTGTACCGACTGTGGAAGGGGAAGCTTT	1350
Db	1554	CCCCGACATTTCAATCACAAGTCTTATGATTAACCTGTACGACATGGAAGGAGATGTTC	1613
QY	1351	CCTAGCAATCTTTGTGCTATGACCACTTACTGAGACCCGACAGCCCTCAGCAAGAG	1410
Db	1614	CCGAGCAATCTGTGCTATCACTCACTGCTGTCTCCAGCAACCCCTTAACAAAGAG	1673
QY	1411	AGTGTGTTGTCTCTGCACAGGCGTGTGCCAAAGACAAATGGGTGTGAACCATGAGAC	1470
Db	1674	AGTGTGTGTGTCTCTGTTCAGGGCGGTGCCAAAGACAAATGGGTGTGTGAACCAAGAGAC	1733
QY	1471	CTGTGTGTGGTGAAGCAGATTGGAACGGGGGAACTTTGGCAGAAAGTTTCAAGCGGACCTG	1530
Db	1734	CTGTGTCTGGGAGAGCAGATCGGAAGGGGAACTTCCGAGAGGTGTTTATGTGGCGCTT	1793
QY	1531	CGAGCCGACAAACCCCTGTGTGAGTCTTGTGTGAGAGACGCTTCCACTGACCTC	1590
Db	1794	CGTGCAGACAAATACCCCGGTGTGTGAAGCTTGTCCGAGAGACACTCCACTGACCTC	1853
QY	1591	AAGGCGAAGTTTCTACAGGAAGCAGAGATCTCTGAAGCAAGTCAAGCAACCCAACTGTG	1656
Db	1854	AAGGCGAAGTTTCTCAAGGAAGCAGAGATCTCTGAAGCAAGTCAACCAACCCCAACTGTG	1913

QY	1651	CGTCTCATGTGGTGTCTGGACCCCAAGAGAGCCCACTTCAATCTGTCATGCAATGAGACTTGTGGAG	1710
Db	1914	CGTCTCATGGGGGTCTGCAACACAAACAGCCCATCTACATCTGTCATGAGACTGGTTCAA	1973
QY	1711	GGGGGGCGACTTCTGACCTTCTCTCCGACGGAGGGGGGCCCGCTTGCGGGTGAAGACTCTG	1770
Db	1974	GGGGGGCGACTTCTCACTTCTCTCGGACAGAGGGAGGCCCGCTTGCGGGTGAAGACTCTG	2033
QY	1771	CTGCAGATGATGGGGGGGANTGCAGCTGCTGGCACTGAGTACCTGGAGAGCAATGCTGCATC	1830
Db	2034	CTGCAGATGATGGGGGGAGCAGACTGCTGGCACTGGAATTACTTGGAAAGCAATGTTTGTATC	2093
QY	1831	CACCGGGGACCTGGGTGCTCGAGACTGCTGCTGTGACAGAGAAATGTCCTTGAAATCAGT	1890
Db	2094	CACAGGGACCTGGGTGCTCGGAACTGCTGCTGTGACAGAGAAATGTCCTGGAATCAGC	2153
QY	1891	GACTTTGGGATATGTCCTCCAGAGAGAAAGCCGATGGGTCTATGTCAGCCTTCAAGGGGCTCTAGA	1956
Db	2154	GACTTTGGGATATGTCCTCCGGAAGAGACTGATGGGATCTATGTCGCGCTGCAAGGGGCTCTAGA	2213
QY	1951	CAAGTCCCCCGTAGATGAGACCGCACTGAGGCCCTTAACTTACCGGCCGCTACTCTCTCCGAA	2010
Db	2214	CAAGTCCCCGTTAAAGTGAAGTCTGCCCCCTGAGGCCCTTAACTTACCGAGCTTACTCTCAGAG	2273
QY	2011	AGCGACGTGTGAGACTTTGGCATCTTGTGCTCTGGAGAACCTTCAAGCTTGGGGGCTTCCCCC	2070
Db	2274	AGTATGTGTGAGACTTTGGCATTTTGTGCTGTGGAGAACCTTCAAGCTTGGGGGCTTCAAGCC	2333
QY	2071	TATCCCAACTCTCAGCAATCAGACAGACACAGGAAATTTGTGTGAGAGAAAGGGGGGCGCTCTGCC	2130
Db	2334	TATCCCAACTCTCAACCAATCAGACAGACACAGGAAATTTGTGTGAGAGAAAGGGGGGCAATGCTTGCCT	2393
QY	2131	TGCCCAGAGCTGTGTCTGATGTCCGCTGTTTCAAGCTCTAGAGCAAGTGTGGGGCTTATAG	2190
Db	2394	TGCCCAGAGCTGTGTCTGATGTCCGCTTTCACAGGCTCTATGAGACAGTGTGGGGCTTATAG	2455
QY	2191	CCTGGGACCGGCGCCAGCTTCAGACCACTTACAGAGAGCTTGAGAGCATTCGAAAGCGG	2250
Db	2454	CCTGGGACGAGACCTTATGTTTCAAGCATCATCTGCAGAGAGCTTCAACAGTATCTCGCAAGCGG	2513
QY	2251	CATCGG 2256	
Db	2514	CATCGG 2519	

```

ACTST 8      RESULT 8
LOCUS       ACCTSI140A
DEFINITION  ACCTSI140A
              Fujinami sarcoma virus temperature sensitive (ts) p140 transforming
              protein RNA, 3' end.
ACCESSION   M14930
VERSION     M14930.1 GI:209688
KEYWORDS    transforming protein.
SOURCE      Fujinami sarcoma virus
            Fujinami sarcoma virus
            Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.
REFERENCE   1 (bases 1 to 2715)
            Chen,L.H., Hatada,E., Wheatley,W. and Lee,W.H.
            Single amino acid substitution, from Glu1025 to Asp, of the fps
            oncogenic protein causes temperature sensitivity in transformation
            and kinase activity
            Virology 155 (1), 106-119 (1986)
JOURNAL     87044080
MEDLINE     28775522
PUBMED      28775522
COMMENT     Original source text: Fujinami sarcoma integrated viral DNA, clone
            pFL-5 from ts FL-15.

FEATURES
    source
        location/Qualifiers
            1..2715
                /organism="Fujinami sarcoma virus"
                /mol_type="genomic RNA"
                /db_xref="taxon:11885"
            <1..2697
                /note="p140 transforming protein"

```



```
/codon_start=1
/protein_id="AAA42403.1"
/db_xref="GI:209689"
/translation="PHDVTPMPEVILGPAPYALIMPMAMASQJLHPPOEHTSTSA
AGTWRLQASBSRHLPHCSAPSHODSHMGFBPLCPKSGISLRLQDSLRLE
LMKRWMSORAKSDREYAGMLHMFSQLERKDEGHLRADHDSWMLSVLASQET
LSQTLRHAELIAPLAKSLILIRDOQLAKVFSOMQWLLSOYVMTTQOAEKKA
QVSLVRSSTOAKRKYOASKDKERKAEKYVSLKYSLYALNSOYVMTTQOAEKKA
HYORALPTLHSLYLODMVTLKEILGEVCSITSLQDEVLITHOEVANVEMIE
PATYRSFVQCHRYDSEVPVATFDSLEBASLEPQLQNELTIBSVQHSLEIE
BELLSARKVASKEQVWELQVLRGBELALSGBEVHLGKRGQIREAQDQLQSLVC
AQAKQADRMLTNKLALSGSDPPALPQEDROASCTDORSVTLTKTNIS
GIFSPRSLPPVPLIPEVQKPLCOQMAYGALPRESVQELIYDSVFLVRESQKOE
YVLSVMDQGRHRIIOADNLYRLDDGPTIPLIDHLOSORPIITRSQVILTRA
VLKDWLVNHEVDLIGRIGRGNGRGEFSGRLADNTPVAVNSCRETLRPELAKRLO
EARIKQCHNPVRLIGVCTQKQPIYIYVWELVQSDPLFSRLSKPRLMKKLITOM
DNAAGMEYIESHCHTIRDLAANRCVTEKNITKISDFPGRSREBGBVASTGMQOI
PVKMTPEALNTGMYSESDVMSFGILMWEAFSLGAVPAMLSNQOTREAIIBGVRLB
PPECBEDVRLMQRCMEYDPHRRPSFGAVHODLIARKHR"
```

ORIGIN

Query Match 45.1%; Score 1018.2; DB 14; Length 2715;

Best Local Similarity 66.2%; Pred. No. 2.2e-167; Matches 1641; Conservative 0; Mismatches 608; Indels 228; Gaps 3;

```
QY 1 ATGGGCTTCTCTTGAGCTGTGACGCCCCAGGAGCCACGGGGTCTTGACGAATTCAG 60
Db 223 ATGGGCTTGGGCGGAGCTGTGTGCTCCAGAGGGGACAGTACGTGCTGGCTGACG 282
QY 61 GAGGCGAGCTTGTCTTACTGAGAGGCGATGAGAAAGTGATGGCCGAGGCTTCAAGAGT 120
Db 283 GACAGCGAGTTGGCGCTCTGAGAGTATGAGAAAGTGATGTCAACGCTGCAAGAGC 342
QY 121 GACAGGAGATGACAGACTGCTTCAACCAATGTCC-----CTGACGACAGT 168
Db 343 GACCGGAGATGACCGGGAGTGTGACCAACATGTTCTCAGCTGAGAAACAGAGAGGC 402
QY 169 GGGGGCCAGAGCGGGGCTATCAGCCCTGACGCCCATCATGCTAGTCTGGGCTGAGATC 228
Db 403 CTGGGACATCTCCGTGCCACCGACCAACAGACCAATCCGGGAGTGTGTGGGTTCTG 462
QY 229 ACCAGCCAACTGAGGGGCTGAGCCGCTTCTGCGGACGACCGACAGAGATCTGAATCA 288
Db 463 GCAAGCCAGACTGAGACACTGAGCCAGACGCTGGCGGCGACGCGGAGAGTGGACGCG 522
QY 289 GGGGCTTGAAGCAAGCTGACCTGCTCATCCGGGAAAGGACGACGAGTCTCCCAAGCTTAC 348
Db 523 GGGCCGCTGGCCAAAGCTGACATATCTGATCCGTGACAAAGCAGCTTCCCAAGCTTTC 582
QY 349 AGCGAGATGAGCAGCTGACGAGAGCTCAACCAAGACCCACAGCCAGACATTGAG 408
Db 583 AGTGAAGATGAGCAGAGCTCAAGCCAGAGATACCCCTGAGACCAACAGCAGAGAGTTGAG 642
QY 409 AAGCTGAAGAGCAGTACCGAGCTCTGGCACGAGACAGTCCCAAGCCAAAGCCAAATAC 468
Db 643 AAGCTGAAGAGCAGTACCGCAGCTGTGTGCTGACAGCAACCCAGGCCAAAGTAC 702
QY 469 CAGGAGCCCAAGAAAGCAAGGACCCGTGACAGGCCCAAGAGTATGTGCCAGGCTG 528
Db 703 CAGGAGCCCAAGAAAGCAAGGACCCGAAAGGCGAAAGAAATATCGTGCAGAGCTG 762
QY 529 TGGAGCTCTTGTCTACCAACCCGCTATGTGCTGGGCGTGGCGGCTGCGAGCTTAC 588
Db 763 TCGAAGCTTATGCTTGGCAACAACATAGTGTGTGCTGTGACAGGCGGCTGCTGTGAC 822
QY 589 CACCAAGACCAACCAAGCTCTGCTGCTGCGGCTGTGCTGCGGTCACTGACAGACTTGCAC 648
Db 823 CACCGCATCATCACTACGCGTGTCTGCTGCAACGATCCCTCTACAGCTTGCAG 882
QY 649 GAGGAGATGCTTGCATCTGAGAGAGATCTCTGAGAGAAATACCTGAGAGATTAAGAGCTTG 708
Db 883 CAGGAGATGCTTCTTGTGTAAGAGATTCCTGGGAGAGTACGTGACATCAACAGCTG 942
```

```
QY 709 GTGACGATGAGATGCTGTCATTCACCGGAGATGCTGACCTGCTCCGCATTCAG 768
Db 943 GTGACGAGAGAGATGCTGTCATTCACCGAGAGGTTGCCACGCCGCTCCAGATGATGCAC 1002
QY 769 CTTGAGGCTGAGTACCAAGGCTTCTGCGACAGTATGAGTCCGACCTTACCTCCACCC 828
Db 1003 CCCGCCACGAGTACGACGCTTGTTCAGTCCACAGGTACGACTCTAGAGTACACACCG 1062
QY 829 TGTGTACGTTGATGAGTCACTGCTTGAAGAGGGTGAACCGGTGAGCTGGGAGCTC 888
Db 1063 GCAGTACCTTTGATGAGAGCTTGTGAGAGAGCAGAGAGCTTGAAGCCAGGAGAGCTG 1122
QY 889 CAGCTGAAGAGCTGACTGTGAGAGAGCTGACACACGCTGACCTCACTGACAGATGAG 948
Db 1123 CAGCTGAATGAGCTGACATCGAGAGCGTGCAGCACTCCCTGACATTCATTTAGAGAGAG 1182
QY 949 CTGGCTGTGAGCCACGAGATGCTTCAAGCGGCGAGAGATGTTACGACGCTGCAACAG 1008
Db 1183 CTGTTGGCCAGACGAGAGGACAGTGAACAGACAGACAGCGGCTGTGGAGCTGACGTT 1242
QY 1009 GAGCTCCGAGATGAGAGAGAGAAACACCCACCCCGGAGCGGGTGCAGCTGGGCAAG 1068
Db 1243 GAGCTGCGGGGAGAGAGAGCTGGCACTGAGCCCTTGGGAGCGGATGACATTTGCTGGCAG 1302
QY 1069 AGCAGATGCTGCAAGAACACTGACAGGGGCTGACAGTACGCTGTGACAGCCAGCCAG 1128
Db 1303 CCGCAGAGGCTGCGGGAGGCCACAGGACAGCTGACAGGCTTGGTCTGTGCTCAGGCCAAG 1362
QY 1129 CTGACAGCCCAAGAGATTTGCTGACAGACCAAGCTGAGACACTGTGGGCCCCGGGAGCC 1188
Db 1363 CTGACAGCCCAAGAGATTTGCTGACCAACCAATGAGAGCTGAGCTTGAAGAGCC 1422
QY 1189 CCGCTGTGCTGCTGCTGAGAGTGAACGCGCACTCCAGCTGCTCTGAGAGCAGAGCGA 1248
Db 1423 CTTCCGCTTGGCTTCCCTGAGAGAGACCGGAGTCAAGCTGCTCAACGATCAAGAGCCG 1482
QY 1249 GAGGGGAGAGAGACACCCACGCTGAGATCTTGAAGACCAATCTCAGGAATCTTCCGC 1308
Db 1483 AGTGGGAGTGA-----CTGCGCTGAAGACATCAAGAACCAATCTCGGCGATCTTCAGT 1536
QY 1309 CCCAAGTTCTCG----- 1320
Db 1537 CCCAGTTCTCGTGCACCTCCGTCCTTCATCCGAGAGTGAAGAGCCGCTGTGC 1596
QY 1321 ----- 1320
Db 1597 CAGCAGGCTGTGACACAGGGGCGCATCCCGGCTCGAGAGTGAAGAGCTGTGAATAC 1656
QY 1657 AGCGAGACTTCTGTGCGGAGAGACAGGGGAAAGAGATGTGCTCAGCTGTGCTG 1716
Db 1321 ----- 1320
QY 1321 -----AACTTATCCGATCTGAA 1338
Db 1717 TGGGATGGCAGGCCCGGACCTTCAATCATTCAGGCTGTGACAACTGTACAGGCTGAA 1776
QY 1339 GGGGAGGCTTTCCTGACATTCCTTGTCTATGACCACTACTGAGACCCAGCAGCC 1398
Db 1777 GAGCATGAGCTTCCACATCCGCTGCTCATGACCACTGTGACAGACCGGCGCC 1836
QY 1399 CTCAACCAAGAGATGATGTGTCTGACAGGCTGTGCGCAAGACCAAGTGGTGTG 1458
Db 1837 ATCAACCCGAGAGAGCGGACATCGTCTGACCAAGCGGTGTCAAGGACAAAGTGGTGTG 1896
QY 1459 AACCATGAGAGCTGTGTGTGGTGTGAGCAGATTTGACCGGGGAACTTTGGCGAAAGTGTTC 1518
Db 1897 AACCATGAGAGCTGTGTGTGGGAGGAGCCATTTGGCGGGGAACTTTGGCGAGGAGTGTTC 1956
QY 1519 AGCGAGCGCTGAGACCGACCAACACCTGTTGGCGGAGTGAAGTCTTGTGAGAGAGCGCTT 1578
Db 1957 AGCGCGCGCTGTGCTGTGACCAACCCCGTGTGCGGTGAATCTCTGCGGGAAACCTTG 2016
QY 1579 CCACTGACCTCAAGGCCAAGTTTCTTACAGAAAGCAGAGATCTTGAAGAGTACAGCCAC 1638
```



```

Db 2017 CCGCCGAGCTGTAAGCCAAAGTTCCTGCGAGGAAGCAGATCTCTAAGCAGTCAACCAT 2076
Qy 1639 CCCAACATGCTGTGCTCATTTGCTGTGACCCCAAGAGCCCATCTACATGTCATG 1698
Db 2077 CCCAACATGCTGTGCTCATTTGCTGTGACCCCAAGAGCCCATCTACATGTCATG 2136
Qy 1699 GAGCTTGTGACAGGGGGCCACTTCTCTGACCTTCTCTGACAGAGGGGGCCCTCTCCG 1758
Db 2137 GAATCTGTGACAGGGAGGAGCTTCTCTGACCTTCTCTGACAGAGGGGGCCCTCTCAAG 2196
Qy 1759 GTGAAGACTCTGTGACAGATGTGAGGGAGATGACAGTGTGACATGAGTACTGTGAGAGC 1818
Db 2197 ATGAGAGACTGATCTGAGATGATGACATGACATGCGCGGCGGATGAGTACTGTGAGAGC 2256
Qy 1819 AAGTGTGACATGACAGCGGAGCTGTGCTGTGAGACTGTGCTGTGACAGAGAGAAATGTC 1878
Db 2257 AAGCACTGATCTGACAGGAGCTGTGCTGTGAGACTGTGCTGTGACAGAGAGAAATGTC 2316
Qy 1879 CTGAAGATGATGATCTTTGGATGTCTCCGAGAGAGAGCCGATGGGCTCTATGACGCTCA 1938
Db 2317 CTGAAGATGATGATCTTTGGATGTCTCCGAGAGAGAGAGATGTGTCTATGCTCTCAAG 2376
Qy 1939 GGGGGCTCTGACAGCAAGTCCCGTGAAGTGAACCGACCTGAGGCGCTTAACCTAGGCGCG 1998
Db 2377 GGGGGCATGAGAGAGATCTCCCGTGAAGTGAACCTGAGGCGCTTAACCTAGGCGCTG 2436
Qy 1999 TACTCTCTGAGAAAGCAGCTGTGAGCTTTGGATCTTGTCTGTGAGAGCTTCAAGCTTG 2058
Db 2437 TACAGCTCGAGAGAGCAGCTGTGAGCTTTGGATCTTGTGTGTGAGAGCTTCAAGCTTG 2496
Qy 2059 GGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2118
Db 2497 GGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2556
Qy 2119 GGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2178
Db 2557 GTGCGGCTGAGAGCGCCCGGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2616
Qy 2179 TGGGCTATGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2238
Db 2617 TGGGAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2676
Qy 2239 ATCCGAAAGCGGATCG 2255
Db 2677 ATCCGAAAGCGGATCG 2693

RESULT 9
ACF 4788 bp ss-RNA linear VRL 27-APR-1993
LOCUS Fujinami sarcoma virus (unintegrated circular), complete genome.
ACCESSION J02194.1 GI:209686
VERSION J02194.1 GI:209686
KEYWORDS C-myc proto-oncogene; complete genome; fps oncogene; polypeptide.
SOURCE Fujinami sarcoma virus
ORGANISM Fujinami sarcoma virus
REFERENCE 1 (bases 1 to 4788)
AUTHORS Shibuya, M. and Hanafusa, H.
TITLE Nucleotide sequence of Fujinami sarcoma virus: evolutionary relationship of its transforming gene with transforming genes of other sarcoma viruses
JOURNAL Cell 30 (3), 787-795 (1982)
MEDLINE 83050964
PUBMED 6291784
REFERENCE 2 (bases 1832 to 1881; 2851 to 2902)
AUTHORS Carlberg, K., Chamberlin, M.E. and Beemon, K.
TITLE The avian sarcoma virus PRC11 Jacks 1020 nucleotides of the fps transforming gene
JOURNAL Virology 135 (1), 157-167 (1984)
MEDLINE 84225820
PUBMED 6328746

```

```

COMMENT Original source text: FSV [Fujinami sarcoma virus], unintegrated
circular DNA [1], [2]
The sequence of 1182 amino acids was deduced from the Fujinami
sarcoma virus (FSV) transforming protein p130, the product of
gag-fps fused gene. p130 is highly homologous to the gag-fes
sequence of feline sarcoma virus (see separate entries), and is 40%
homologous in the 280 residue carboxyl end (including the
phosphoacceptor tyrosine residue) of Rous sarcoma virus p60
protein.

FEATURES
source
    1..4788
        /organism="Fujinami sarcoma virus"
        /mol_type="genomic RNA"
        /db_xref="taxon:11885"
    LTR
        1..347
            /note="5' LTR"
        repeat_region
            247..267
                /note="5' terminal repeat"
        CDS
            626..4174
                /note="p130 polypeptide"
                /codon_start=1
                /protein_id="AAA42402.1"
                /translation="MEAVIKVISAQKTYCGKTSPEKKEIGAMLSOLQKELMSLSD
                LVSQSGWBPITATLQORAMVLKSGELKTWGLVLAALKAREOYTSBAKMTLIGG
                GVSPPGPECTEKPRATERIDKGETTVQDITQAPERTPTPTVGTSCHCGRALICN
                CATASAPPPYVSGSLVPSLAVGBOGGGSGTTPKAPBAPBAPBAPBAPBAPBAPBAPB
                ARIARELASTGPPMAMPVAVITGEBAGMPLPKIAGLAVGAGAGLAVAVAVE
                ALMSSEPLDHPVAVITGPAALPMMDANAASQGLRPOPEBTSTSAAGTARL
                TOASESRHLPHCSAAPSODHSAMGFGELPKQKHTLRLQDESLRLLEMKMM
                SORAKSDREYALHMHMPSOLKOEGLHRLARSDSOIGESMWVLAOTETLSOTLR
                RAABELAARPLAKSLILDKOOLKRVSEBQOQSBOYAWTQOQVEBTLKQYRSLV
                RSTQAKRTQASDKEREKAKERVSLKVLALVAVLAVQAAALHHHHTQRA
                LPTLHSLVLSQEEVVLVKEILGEVCSITSLVQEBVLAIHQVAAVEMIDATYS
                SEVOCHRYDSEVPVATPDESILBEAENPEPELQNELTISVOHSLTISIBELLAS
                RKAWSKEORVMELOVELRGEELALSPGRVHLGRQGLREAOQLOGLVCAQAKLO
                AORDMANKLAEIASEPPEPALPIODRQASSTOERSGVTAIKTIKHHSIGIFSR
                PSLLPVPVPIIPVQKPLCOQANYHGAIPRSVQELKTYGDPYRSQKQEVLSVL
                WQGPVHFLITQADNLYLEDGLPTLPIHLQDSQPIITRKSIVLIRATLKRM
                VNHEDVHLGERIIGRNFGEVFSGRIRANTVPAVVSCTETLPBKLAFLOARLIK
                OCNHPIVRLIVCTQKQPIYIVMEVQGDPLVFSRSGPLKMKKLKMMENAAAG
                MEYLSKCHIHDLARNCLVTEKNTKLSIDFGMSQEDBDGYVASTGKMKQIPVKTAT
                PALNAGWYSSESDVSGIILMEAPISGLAVPYANLSNQTREAIHQVRLPEPCP
                EDVYRLMQRCHETDPRRPSFGAHODLIAIKRRH"
            mat_peptide
                629..1549
                    /product="gag protein"
            mat_peptide
                1550..4171
                    /product="v-fps transforming protein"
            LTR
                4442..4788
                    /note="3' LTR"
            repeat_region
                4688..4708
                    /note="3' terminal repeat"
            ORIGIN 5' end of proviral genome; 768 bp upstream of SmaI site.

Query Match 45.1%; Score 1018.2; DB 14; Length 4788;
Best Local Similarity 66.2%; Pred. No. 2e-167;
Matches 1641; Conservative 0; Mismatches 608; Indels 228; Gaps 3;

```

QY	225	ACGAGCCAAATGAGGGGCTGAGGCGCTTGCTGGCGGACGACGAGAGATCTGAACTCA	288
Db	1940	GCAAGCCAGACTGTAAGACATGAGCGACAGCGCTGGCGGGACGCGGAGAGACTGGACCG	1999
QY	289	GGGCGCCCTGAGCAAGCTGAGCGCTGCTCATCCGGGAAAGGGACAGCTTCCGAAAGCTTAC	348
Db	2000	GGGCGCGCTGGCCAAAGCTGAGCATATGATCTGTTGAAAGAGGACGAGCTTCCGAAAGTCTTC	2059
QY	349	AGCGAGCATGTGGCAGCAGCTGCAAGAGACTCACCAAAGCCACAGCCAGGACATTTGAG	408
Db	2060	AGTAGGCAGTGGCAGCAGACTCAGCGCAGGAGATGCCCTGGAGCACACAGCAGAAAGTTGAG	2119
QY	409	AAGCTGAAGAGCCAGTATCCGAGCTCTGGCACGGGACAGTGTGCCAAAGCCAAAGCGAAATGAC	468
Db	2120	AAGCTGAAGAGCAGATATCCGAGCGCTGTGCTGATGACAGACCCAGGCGCAAGGATGAC	2179
QY	469	CAGAGGCGCAGCAAAAGACAAAGGACCGTGACAAAGGCCAAAGACAAATATGTGCGAGCTG	528
Db	2180	CAGAGGCGCAGCAAAAGACAAAGAGGCGGGAAAAAGCGAAAGTACGTGCGAGCCTG	2239
QY	529	TGGAAAGCTCTTTGCTCAGCAACAACCGCTATGTGCTGGCGCGGCGGGCTGCGCACTTAC	588
Db	2240	TCGAAGCTCTATGCTTGGCACAACCAAGTACGTGCTGTGCAAGGCGGCGCTTGGAC	2299
QY	589	CACCAAGCCACAACAAGCAGCTCTGTGCGCGGCTGCTGCGGTGACTGACAGACCTTGAC	648
Db	2300	CACCAACCATCATACAGCGGTGCTGTGCGCAACCTGACAGAGTCCCTCTACAGCTTGAC	2359
QY	649	GAGAGATATGCTTGATCTCTGAAGAAATCTCTGACAGAAATCTGAGAAATTAAGACCTG	708
Db	2360	CAGAGAGATGATCTTGTGTTTAAAGAGATTTTGGGAGATCTAGACATCAAGTCTG	2419
QY	709	GTGACAGATGAGTGTGTGGCCATTCAACGGGAGATGCTGAGCTGCGCCGATCCAG	768
Db	2420	GTGACAGAGAGACGTGTGCGCATCAACGAAGGTGTGCCAGCGCGTGCAGATGATGAC	2479
QY	769	CTGAGGCTGAGTACCAAGGCTTCTGTGCAAGTATGAGGTCCGACCTGACGTCCACCC	828
Db	2480	CCCCGCACTGATGTCAGCAGGCTTTGTTGATGTCACAGGTACGACTGAGATGACACCG	2539
QY	829	TGTGTCAACGTCGATGATGATCACTGCTTGAAGAGGGTGAACCGCTGAGGCTGGGGAGCTC	888
Db	2540	GCAGTGAACCTTTGATGATGAGACTGTGCTGAGGAGGCGAGAACTTGAACTCAGGGGAGCTG	2599
QY	889	CAGCTGAACGAGCTGACTGTGAGAGCGGTGACGACACGCTGACCTCAGTGAAGATGAG	948
Db	2600	CAGCTGAATGAGCTGACCATCGAGAGCGTGCMACTCTCCTGATCATTTGAGGAGGAC	2659
QY	949	CTGAGCTGTGGCCACCGAATATGTGTTCAAGGGGGCAGAGATGTGTTAGCCACTGCAACAG	1008
Db	2660	CTGTTTGGCAGACGAGAGAGGACAGTGAAGCAAGAGGACGCGGATGTGGAACTGTGAAGTT	2719
QY	1009	GAGCTCCGGAATGAAGAGGAGAAACCCACCCCGGAGCGGGGTGACACTCTGGGCAAG	1068
Db	2720	GAGCTCGGGGGCGAGGAGACTGGCACTGACCCCTGGGAGGCGGGTGCATTGTCTGGGCAAG	2779
QY	1069	AAGCAAGTGTCTGCAAGAGACACTGACGAGGGCTGCAAGTATGCGCTGTGACCCAGGCCAG	1128
Db	2780	CGGCGAGGGGCTGCGGGAGGCCACAGACAGGCTGCAAGGCTTGTGTCTGTCAAGGCCAAG	2839
QY	1129	CTGCAAGGCGCAGCGAGGATGTGCTGACGCAAGCTGTGAGCACTGTGGGCGCGGGGAGGCC	1188
Db	2840	CTGCAAGGCGCAGCGGGATATGTGCTGCAACAAGCTGGGCAAGCTGGGCTTGTAGAGAGCCC	2899
QY	1189	CCGCTGTGTCTGCTCTCTGACAGATGACCGCACTTCAAGTGTCTCTCGAGACAGAGGCA	1248
Db	2900	CTGTCCGCGCTTGCCCTGACGAGAGGACCGGAGAGTCAAGGCGCGCTCACAGGATCAAGAGCGC	2959
QY	1249	GAGGGGGGAGAGCAACCAAGCTGGAATCTTTAAGACCAATTTCAAGAAATCTTCCG	1308
Db	2960	AGTGGGGTGA-----CTGCGCTGAAGACATCAAGAAACACATCTCGGGATCTTCAGT	3013

QY	1309	CCCAAGTTCTGG-----	1320
Db	3014	CCCAAGTTCTGGCTGCGCCACCTCCGTCGCCCTCATTCGCCAGAGTGCAGAAAGCCGCTGTGC	3073
QY	1321	-----	1320
Db	3074	CAGCAGGCTGGTACCACGGGGCCATCCCGCGCTCGAGAGTGCAGGAGGCTGGAATAC	3133
QY	1321	-----	1320
Db	3124	AGCGGAGACTTCTGTGTCGGGAGAGACGAGGGAAAGCAGAGTATGTGCTCAGCGTGTG	3193
QY	1321	-----AACTGTATCCGACTGGAA	1338
Db	3194	TGGATGGGAGACCCCGGCACTTCATCATCAGCGTCTGTGACAAACTGTATCAGGCTGGAA	3253
QY	1339	GGGGAAGGCTTTCCTAGCATTCCTTGTGCTCATCGACCACTACTGTAGACCCAGACAGCC	1398
Db	3254	GACGATGGCCTTCCACCACTCCCGCTGCTCATTCGACCACTGCTGCAGAGCCAGCGGCC	3313
QY	1339	CTCACCAAGAAAGTGTATTGTCTCTGCA CAGGGCTGTGCCCAAGACAAGTGGTGTG	1458
Db	3314	ATCACCCGCAAGAGCGGATGTCTCTGAC CAGAGCGTGTCTCAAGACAAGTGGTGTG	3373
QY	1459	AACTGTAGGACCTGTGTGTGGGTGACGAGATTGGACGCGGGAACTTTGGCCAAAGTCTTC	1518
Db	3374	AACTGTAGGACGTCGTGTGTGGGGAGCCCATTTGGCCGGGAACTTCGGGAGGTGTTC	3433
QY	1519	AGCGGAGCGCTGCGAGCCGACAA CACCCTGTGTGGCGGTGAAGTCTTGCGAGAGACGCTC	1578
Db	3434	AGCGGCGGCTGCGTGTGTGACAA CACCCTGTGTGGCGGTGAAGTCTTGCGGGAACCTTG	3493
QY	1579	CCACCTGACCTTCAGAGCCCAAGTTTCTACAGAA GCGAGATCTGTGACAGTACAGCCAC	1638
Db	3494	CCGCCGAGCTGAGAGGCCAA GTTCTGTGCGAGAA GCGAGATCTCTCAAGCAGTCAACAT	3553
QY	1639	CCCAACATGTGTGCTGTCTATTGTGTGTGCA CCCCAGAACAGCCCATCTATCATGTCTATG	1698
Db	3554	CCCAACATGTGTGCGGCTCATGTGGCGTGTGCA CCCCAGAACAGCCCATCTATCATGTCTATG	3613
QY	1699	GAGCTGTGCAAGGGGGGCGACTTCCGTGACCTTCCTGCGACGGAAGGGGGGCGCGCTGCGG	1758
Db	3614	GAGCTGTGTGCAAGGGAGGGGACTTCTGTAGCTTCTGTGCGACAGCAAGGGGCCCGCTGTAG	3673
QY	1759	GTTGAAGACTGTGCTGCAGATGTGTGGGGGATG CAGCTGTGTGCAATGAGTACCTGTGAGAGC	1818
Db	3674	ATGAAGAAGCTGATCATGAATGATGAGAA TGC CGCGGGGCAATGAGTATCTGTGAAGC	3723
QY	1819	AAGTGCTGATCATCCCGGAGCTGGCTGCTCGGA ACTGCTGTGTGACAGAGAAAGATGTG	1878
Db	3724	AAAGCATGTGATCACAAGGAGCTGTGGTGTGCC CACATGCTGTGTGACAGAAAAGAACCC	3793
QY	1879	CTGAAGATCAGGACCTTGTGGATGTGTC CGAGAGAAAGCCGATGGGGGTCTATGACGCTCA	1938
Db	3794	CTGAATAATCAGGACTTGTGGATGTGTC CGGACAGAGAGAGATGTGTCTATGTCTCCACG	3853
QY	1939	GGGGGCTCAGACAAGTCCCGCTGAA GTGAGACCGCACTTGAGGCCCTTTAACTACGCGCG	1998
Db	3854	GGGGGCTATGAACAGATATCCCGGTGAA TGTGACTGTGCCCGGAGGCTGTGAATTAACGCTGG	3913
QY	1999	TACTCTCTCGAAAAGCAGCGTGTGAGCTTTGGCAT TTTGCTCTGGAGACCTTCAAGCTTG	2058
Db	3914	TACACTCTCGAAGAGCGACGTGTGAGCTTTGGGATCTGTGTGGGAAAGCTTCAAGCTTG	3973
QY	2059	GGGGGCTCCCTTATTCCTCAACTTCAGCAATTCAGCAACA CCGGAGTTTGTGTGGAAGGGG	2118
Db	3974	GGCGCGGTGCTCTTACGCCCAACTCAGCAACA CAGCAACA CCGGCGAGGCTCATCGACAGAGGC	4033
QY	2119	GGCGGTGCTGCTGCGCCAGAGCTGTGTCTGTATGCGGTGTTCAAGGCTCATGTGAGCACTGTG	2178
Db	4034	GTGCGGCTGTGAGCCCGCGAGCAGTGTCCCGCAGAGAGTGTATCCGCTGTATGTGACGCGGTGC	4093
QY	2179	TGGGCGTATGAGCTGTGGAGCGGGCCCAAGCTTTCAGCA CCACTTACAGAGACTGTGCAAGC	2238

Db 4094 TGGAGTACACCTTCACAGCGCGCGAGCTTGCGCCCTCCACGAGACTTCATCGCC 4153
Qy 2239 ATCCGAAGCGCATCG 2255
Db 4154 ATCCGCAACGCGACCG 4170

RESULT 10
AF033810 4788 bp RNA linear VRL 03-DEC-1998
LOCUS Fujinami sarcoma virus, complete genome.
DEFINITION AF033810
ACCESSION AF033810
VERSION AF033810.1 GI:2801466
KEYWORDS
SOURCE Fujinami sarcoma virus
ORGANISM Fujinami sarcoma virus
REFERENCE 1 (bases 1 to 4788)
Retrovirus; Retroviral taxonomy; Retroviridae; Alpharetrovirus.
AUTHORS Retrovirus; C.J.
TITLE Appendix 2: Retroviral taxonomy, protein structure, sequences, and genetic maps
(in) Coffin, J.M. (Ed.);
RETROVIRUSES: 757;
Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York,
NY, USA (1997)
2 (bases 1 to 4788)
Chapman, C.
Direct Submission
Submitted (12-NOV-1997) NIH, NLM, Rockville Pike, Bethesda, MD
20894, USA

FEATURES
source
1..4788 location/Qualifiers
/organism="Fujinami sarcoma virus"
/mol_type="genomic RNA"
/db_xref="taxon:11885"
22..4442
/product="p140 polyprotein"
22..102
380..3928
/codon_start=1
/product="p140 polyprotein"
/protein_id="AAC82565.1"
/db_xref="GI:2801467"
/translation="MEAVIKVISACKTTCGKTSPEKKEIGAMLSQLQREGLIMLSLD
LVSFGMDPTTALTORAWYIGSGELKTWGLVGLAKAREQVTSBOAKFWLGAG
GAVSPGPECIEPATERIDKGETTVORDTKMAPEITAPKTVGSCYHCAIGCN
CATASAPPPYVSGSLYPSLAGVBOGQGDTPRGAEOBARGRGTGALPALTDM
ARTRELAGSPMVAMPVITKTEGPAWTPLEBKLAGLAGAGAGAGSLPRAVAVE
ALMSRPLPHDVTNPMRVILGRPHALMDMAASQOLRPOQERTSTSAAGTTRL
TQASERHRLPHCSAAPSODHSAMGEGPELWCPKHTELRQLDSELRLEIMKTM
SOAKSDREYAGMLHMFSQLKEQEGELGHLRATDHSQIGESVWVLAQETLSQLR
RHAEBLAGLAKLSILIRDKQQLRKYFSQMOQLSOEYAWTQOEVEKKAQYRSLV
RDTQAKRKYOBSKDKEREKAKYRSLKLYALNOVLAVOAALHHHHYORA
LPTHELSVLSQEMVLVLEILGEYCSITSIVODVLAHQVLAHVMIDPATYS
SPQCHRYSSEVPATVPBESLLEBAMNEPBELOMELTIESVQSLSTIEBELLAS
RKAVSSKEQWELQVLRGEELALSPEEVHLLGRQGLREAOQLQGLVCAQALQ
AQRDMLANKLAELSESEPPALPQEDROGARSITDERSVTALKTKHISGIFSR
FSLPVPVPLPEVQKPCQAWYHGAIPREVOELKYSQDFLRESQOEVLSTVL
MDQPRFTIIOADNLYRLEDDGLPTIPLIDHLSORPTKSGIVLTRAVALKDKM
VNHEDVLGERITGRNGPGEVSGRLRANTPVAVSCRETPLELKATLOARILK
OCHRNIVRLIGCTOKOPYIYIMELVOGDFLSPRSKQPRILKMKLLKMBENLAG
MEYLESKCHIRDLAARNCILVTEKNTLKTSDFGMSQOEDEGVASTGEMKQIPVKATA
PEALNGWSSSESDVMSFGILWEAFSLGAVFANLSNOOTRAIEQVRLPEQCP
EDVYRLMORCMEYDPHRRPSFGAVHODLAIIRKRRH"

mat_peptide
899..1087
/product="p14 MA"
mat_peptide
899..1087
/product="p10"
mat_peptide
1085..1309
/product="CA"
mat_peptide
1307..13925
/product="FBS"

3'UTR 4196..4442
ORIGIN
Query Match 45.1%; Score 1018.2; DB 14; Length 4788;
Best Local Similarity 66.2%; Pred. No. 2e-167;
Matches 1641; Conservative 0; Mismatches 608; Indels 228; Gaps 3;
Qy 1 ATGGGCTTCTCTTTCAGCTGTGACAGCCCGCCAGAGGCGACAGGGTCTCTGACGAAATGAC 60
Db 1454 ATGGGCTTTCAGGCGGAGCTGTGTGTCCGGAAGGCGGACACCTGAGCTGTGCGGCTGAC 1513
Qy 61 GAGGCGAGCTTGTCTTACTGTGAGGCGGACGAGAAATGTGATGCGGCGGCTCAAGAGT 120
Db 1514 GACAGCGAGTGTGGCTCTCTGAGCTGTGAGAAATGTGATGCTGAGCTGTGAGAGC 1573
Qy 121 GACAGGAGTATGACAGAGCTCTTCAACCATGTCC-----CTGACGACAGT 168
Db 1574 GACCGGAGTATGACGCGGAGTGTGACACCATGTCTTCAAGCTGTGAGAAACAGAGGCG 1633
Qy 169 GGGGGCCAGAGCGGGCCATCAGCCCTGACAGCCCATCATGATCTTGGCTGAGATC 228
Db 1634 CTGGGACATCTCCGTCCACCGACACAGACAGCATCGGGAGTCTGTGGGTCTTG 1693
Qy 229 ACCAGCCAACTGAGGGCGCTGAGCCGCTGTGCGGACAGACGAGATCTGAATCA 288
Db 1694 GCAAGCCAGACTGACATGACCTGAGCCAGAGCTGTGCGGCGGACGAGAGCTGCGACGG 1753
Qy 289 GGGCGCCCTGAGACAGCTGAGCTCTCATTCGGGAAAGCGAGAGCTTGGCAAGCTTAC 348
Db 1754 GGGCGCCCTGAGACAGCTGAGCTCTCATTCGGGAAAGCGAGAGCTTGGCAAGCTTAC 1813
Qy 349 AGCGAGAGTGTGAGAGCTGACAGCGAGAGCTTACCAAGACCAAGAGCAATTGAG 408
Db 1814 AGTGAAGAGTGTGAGAGCTGACAGCGAGAGTACGCTTGAACCAAGAGCAAGTGTGAG 1873
Qy 409 AAGCTGAAGAGCTGATCCAGAGCTTGTGACAGGAGCAGTGCAGCAAGCCAGCAAGTAC 468
Db 1874 AAGCTGAAGAGCTGATCCAGAGCTTGTGACAGCGAGAGTACGCTTGAACCAAGAGCTGAC 1933
Qy 469 CAGAGGCGCAGCAAGAGAGAGCGGTGACAGGAGCAAGAGCAAGTATGTGCGAGCTG 528
Db 1934 CAGAGGCGCAGCAAGAGAGAGCGGTGACAGGAGCAAGAGCAAGTATGTGCGAGCTG 1993
Qy 529 TGGAGCTCTTGTCTACACCAACCGCTATGTGTGTGCGGTGTGCGGCTGCGAC 588
Db 1994 TCGAAGCTCTATGCTCTTGAACCAACCGCTATGTGTGTGCGGTGTGCGGCTGCGAC 2053
Qy 589 CACGAGCAGCAGCAGCAGCTCTGCGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCG 648
Db 2054 CACGAGCAGCAGCAGCAGCTCTGCGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCG 2113
Qy 649 GAGAGAGTGTGCTGATCTCTGAGAGAGATCTGAGAGAAATCTGAGATTAAGTACCTG 708
Db 2114 GAGAGAGTGTGCTGATCTCTGAGAGAGATCTGAGAGAAATCTGAGATTAAGTACCTG 2173
Qy 709 GTGACAGAGTGTGCTGATCTCTGAGAGAGATCTGAGAGAAATCTGAGATTAAGTACCTG 768
Db 2174 GTGACAGAGTGTGCTGATCTCTGAGAGAGATCTGAGAGAAATCTGAGATTAAGTACCTG 2233
Qy 769 CCGAGAGTGTGATGACAGAGCTCTGCGACAGTATGAGTGTGCGGCTGTGCGGCTGTGCG 828
Db 2234 CCGAGAGTGTGATGACAGAGCTCTGCGACAGTATGAGTGTGCGGCTGTGCGGCTGTGCG 888
Qy 829 TGTGTACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
Db 2294 GAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2353
Qy 889 CAGCTGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
Db 2354 CAGCTGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2413
Qy 949 CTGGCTGTGCGCAGAGTGTGCTGAGAGGCGAGAGTGTGCTGAGAGGCGAGAGTGTGCTG 1008

Db 2414 CTGTTGGCCAGCAGAAAGGCGATGAGCAGCAAGAGCAGCGGCTGTGGAGCTGACGATT 2473
QY 1009 GAGCTCCGGAAATGAAGAGAGAAACCCACCCCGGAGAGGGGTGACGCTGTGGCAAG 1068
Db 2474 GAGCTGGGGGCGAGAGCTGGGCACTAGCCCTGGGGAGGGGTGCACTTGTGGGCAAG 2533
QY 1069 AGGCAAGTGTGCGAAGAAAGCATGCAAGGAGCTGCAAGTACGCTGTGACAGCCAGGCAAG 1128
Db 2534 CGGCAAGGGGCTGGGGAGGCGCCAGCAGACGCTCAAGGGCTTGTGTCTGAGGCAAG 2593
QY 1129 CTGCAAGCCCAAGAGATTGCTGCAAGCCAAAGCTGAGACACTGGGCCCCGAGAGCC 1188
Db 2594 CTGCAAGCCCAAGAGATTGCTGCAAGCCAAAGCTGAGACACTGGGCTTGTGAGAGCC 2653
QY 1189 CGGCTGTGTGCTGCTGCGAGATGACCCGACCTCCAGCTGCTGCTGAGAGAGAGCA 1248
Db 2654 CTGTCGCGCTTGGCCCTGCAAGAGAGCCGAGCTCAGCGCTTCACAGATCAGAGAGCC 2713
QY 1249 GAGGGGGGAGAGACACCCAGCTGAGAGATCTTAAAGCCACATCTCAGGAATCTTCCGC 1308
Db 2714 AGTGGGGTGA-----CTGGCTGAAGAGACATCAAGAACCACTTCGGGCATCTTCAGT 2767
QY 1309 CCCAGTTCTCG----- 1320
Db 2768 CCCAGGTTCTCGCTGCCACCTCCGCTGCTCATCCGAGAGGTGAGAGCCGCTGTGC 2827
QY 1321 ----- 1320
Db 2828 CAGCAGGCTGTGATCCAGGGGGCCATCCGCGCTCGAGAGTGCAGAGAGCTGTGAATAC 2887
QY 1321 ----- 1320
Db 2888 AGCGGAGACTTCTGCGTGGCGGAGAGCCAGGGAGAGAGATGATGCTCAGCGTGTG 2947
QY 1321 -----AACTGTATCCAGCTGGA 1338
Db 2948 TGGAGTGGGAGCGCCCGGCACTTATCATCAGAGCTGTGACAACTGTGACAGCTGGAA 3007
QY 1339 GGGGAAAGGCTTCTTATGATCTTCTGATGACCACTTACTGAGACACCCAGCAGCC 1398
Db 3008 GACGATGAGCTTCTCCACCATCCGCTCTCATGACCACTGCTGAGAGACCCAGCGGCC 3067
QY 1399 CTGACCAAGAAAGTGTGTGTCTGTCAGAGGGCTGTGCCAAAGCAAGTGGGTCTG 1458
Db 3068 ATCAACCCGAAAGCGGCACTGCTCTGACAGAGCGTGTCAAGAGCAAGTGGGTCTG 3127
QY 1459 AACCTAGAGAGACTGCTGTGTGTGTGTGTCAGAGGGCTGTGCCAAAGCAAGTGTTC 1518
Db 3128 AACCTAGAGAGACTGCTGTGTGTGTGTGTCAGAGGGCTGTGCCAAAGCAAGTGTTC 3187
QY 1519 AGCGGAGCGCTGCGAGCGCAACAACCTGTGTGTGTGTGTCAGAGAGCGCTC 1578
Db 3188 AGCGGAGCGCTGCTGT 3247
QY 1579 CCACTGACCTCAAGGCGCAAGTTCTTCAAGAAAGCAAGATCTTGAAGCAAGTAC 1638
Db 3248 CCGCGGAGCTGAAGGCGCAAGTTCTTCAAGAAAGCAAGATCTTGAAGCAAGTAC 3307
QY 1639 CCCAACTGT 1698
Db 3308 CCCAACTGT 3367
QY 1699 GAGCTTGTGAGGGGGGCGACTTCTGACCTTCTGCGCAGGAGGGGGCGCGCTGCGG 1758
Db 3368 GAGCTTGTGAGGGGGGCGACTTCTGACCTTCTGCGCAGGAGGGGGCGCGCTTGAAG 3427
QY 1759 GTGAAGACTGTGCTGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1818
Db 3428 ATGAAGAGCTGATCAAGATGATGAGAAATGCGCGGCGGCAATGAGTACTGTGAAGC 3487
QY 1819 AAGTGTGATCCACCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1878
Db 3488 AAGCACTGATCCACGAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3547

QY 1879 CTGAAGATCATGATCTTTGGATGTGTCCGAGAGAGAGCCGATGGGCTTATGACGCTCA 1938
Db 3548 CTGAAGATCATGATCTTTGGATGTGTCCGAGAGAGAGATGTGTATGCTCTCACG 3607
QY 1939 GGGGGCTCAAGCAAGATCCCGTGAATGATGACCGACCTGAGAGCCCTTAACTACGGCCG 1998
Db 3608 GGGGGCTCAAGCAAGATCCCGTGAATGATGACCGACCGAGGCTTGAATTAAGGCTGG 3667
QY 1999 TACTCTCCGAAAGGAGCGTGTGAGCTTTGGATCTTGTCTGTGGAGACCTTCAAGCTG 2058
Db 3668 TACAGCTCCGAAAGGAGCGATGTGTGAGCTTTGGATCTTGTGTGGAGACCTTCAAGCTG 3727
QY 2059 GGGGCTTCCCGTATCCCACTCAAGATCAAGACAGACAGGAGTTTGTGAAGAGGG 2118
Db 3728 GGGGCGTGCCTTCAAGCAAGCTTCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3787
QY 2119 GGGGCTGTGCTTCCGAGAGCTGTGTCTGATGCTGTGTGTGTGTGTGTGTGTGTGTGT 2178
Db 3788 GTGCGGCTGAGAGCCCGGAGAGAGTGTCCCGAGAGAGTGTATCGCTGTATGACAGCTGC 3847
QY 2179 TGGGCTTATGAGCTTGGGAGAGGAGCCAGCTTCAAGACCATTAAGAGAGTGTGAGAG 2238
Db 3848 TGGAGTACAGACTTCAAGAGGAGCCGAGCTTGGGCGCTTCCAGAGAGCTTATGCGC 3907
QY 2239 ATCCGAAGCGGATCG 2255
Db 3908 ATCCGAAGCGGATCG 3924

RESULT 11
M22820
LOCUS M22820 4901 bp DNA linear UNA 04-AUG-1993
DEFINITION Figure 4. Nucleotide sequence of the proviral DNA of the variant clone of ST-FesV.
ACCESSION M22820.1 GI:340749
VERSION M22820.1 GI:340749
KEYWORDS
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 4901)
AUTHORS Kemp,A.W., Bloemers,H.P. and Van de Ven,W.J.
TITLE Structural analysis of a variant clone of Snyder-Theilen feline sarcoma virus
JOURNAL Virus Res. 8 (4), 349-361 (1987)
MEDLINE 88129048
PUBMED 2829460

FEATURES
source Location/Qualifiers
1..4901
/organism="unclassified"
/mol_type="unassigned DNA"
/db_xref="taxon:12908"

ORIGIN
Query Match 37.3%; Score 841.4; DB 13; Length 4901;
Best Local Similarity 78.3%; Pred. No. 1,2e-136;
Matches 1122; Conservative 0; Mismatches 101; Indels 210; Gaps 1;
QY 1033 ACCCAACCCCGGAGGCGGTGACGCTGTGGGAGAGGCAAGTGTGCAAGAGCACTG 1092
Db 1874 ACCCAACCCCGGAGGAGGAGCTGTGGCCAAAGAGAGGTGTGCAAGAGGCGCTG 1933
QY 1093 CAGGGCTGAGGTACCGCTGTGTGACCGCCAACTGTGAGGCCCAAGAGGAGGCTGTG 1152
Db 1934 CAGGGCTGAGGTACCGCTGTGTGACCGCCAACTGTGAGGCCCAAGAGGAGGCTGTG 1993
QY 1153 CAGACCAAGCTGAGACCTGGGCGCGGAGAGCCCGGCTGTGTGTGTGTGTGTGTGTGTGT 1212
Db 1994 CAGGCAAGCTGAGAGAGCTGTGGGCGCGGAGAGCCCGGCTGTGTGTGTGTGTGTGTGTGT 2053
QY 1213 GACCGCACTCAAGT 1272

D	b		2054	GACCGCACTCACAAGTGTCTCCGGAGAGGAGGAAAGGGGAAGACACCACCTTG	2113
Oy		1273	GAGATCTCTTAAGAGCCACATCTTCAGGAATCTTCCGCCCAAGTTCTCG----	1320	
D	b	2114	GAGATCTCTTAAGAGCCACATCTCAGGAATCTTCCGCCCAAGTTCTCGCTCCCTCACCC	2173	
Oy		1321	-----	1320	
D	b	2174	CTGCAGCTCGTACAGAGGTGCAGAAGCCCCTGACGAGACTGTGTGACCAAGGGGCC	2233	
Oy		1321	-----	1320	
D	b	2234	CTCCAGAGGGAGAGGTGGCTGAGCTGTGACGACTCTGGGACTTTCGTGTGCGGAG	2293	
Oy		1321	-----	1320	
D	b	2294	AGCCAGGGCAGAGCAATAATGTGCTGTGCTGTGGAGCGGCCAGCCCCGACATTTC	2353	
Oy		1321	-----AACCTGATCCGATCTTGGAAGGGGAAGGCTTTCCTAAGCATTCCT	1362	
D	b	2354	ATCATCCAGTCCGCTGACMACTCTTACCGATGGAAAGAGGTGGCTTTCAGAGATCCC	2413	
Oy		1363	TTCCTCATTCGACACACTTATGAGCACCAGAGCCCCCTCACCAAGAAAGTGTGTTC	1422	
D	b	2414	TTCCTCGTGCACACCTGTGCGCTCCGAGAGCCCCCTCACCAAGAAAGCGGTATTTGTC	2473	
Oy		1423	CTGCAGAGGGCTGTGCCCAAGACAGAGTGGGTGTAACATGAGGACTTGTGTGGGT	1482	
D	b	2474	CTCAACAGGGCTGTGCCCAAGACAGAGTGGGTGTAACACAGAGACTGTGTGTGGGT	2533	
Oy		1483	GAGCAGATTTGGAAGGGGGGAATCTTTGGCGAATGTTCAAGCGACGCTTSCAGAGCGACAC	1542	
D	b	2534	GAGCAGATTCGGGCGGGGGAACTTTTGGAAAGTGTTCAGTGAACGCTGAGGGCGGACAAAC	2593	
Oy		1543	ACCCTGGGGGGGTGAAGTCTTGTGAGAGAGCGTCCCACTGACCTGAAGCCCAAGTTT	1602	
D	b	2594	ACTTTAGTGGCGGTGAATCTTGTGCGAGACACTCCACTGACATTAAGGCCAAGTTT	2653	
Oy		1603	CTACAGGAAGCGAGGATCTCTGAGCAGTACAGCACCCCAACATCTGTCGTCTCATTTGGT	1662	
D	b	2654	CTTCAGGAAGCAAAGATCTGTGAAGCAGTACAGCCACCCCAACATCTGTGCTCATTCGGC	2713	
Oy		1663	GTCCTGCAACCCAGAGAGAGCCCATCTTACATGTCATGAGCTTGTGCAAGGGGGCGACTTC	1722	
D	b	2714	GTCGTGCAACCCAGAGAGAGCCCATCTTACATGTCATGAGCTGTGAGAGGGGGCGACTTC	2773	
Oy		1723	CTGACCTTTCCTCCGCAAGGAGGGGGCCCGCTGGGGGTGAAGACTCTGTGTGCAATGGTG	1782	
D	b	2774	CTGACCTTTCCTGAGAGAGGAGGGAGCCCGCTGGGGAAGAAAGCTGTCTCAATGGTG	2833	
Oy		1783	GGGATACAGCTGTGCGATGGAATTACTTGAGAGCAAGTCTGCATCAACGGGACCTG	1842	
D	b	2834	GGCGACCGCGCGGGGCAATGGAATTACTTGAGAGCAAGTCTGCATCAACGGGACCTTA	2893	
Oy		1843	GCTGCTCGGAACTGCTGTGTGACAGAGAAATGTCTGAAGATCAGTCACTTTGGAGTG	1902	
D	b	2894	GCTGCTCGGAACTGCTGTGTGACAGAGAAAGTCTGAAGATCAGTCACTTTGGAGTG	2953	
Oy		1903	TCCCGAAGGAAAGCCGATGGGGTCTATGTGACACTCAAGGGGGCTTCAGCAAGTCCCCGTG	1962	
D	b	2954	TCCCGAAGGAAAGCCGATGGGGTCTTACCGGGCTTCAGAGGGGGCTTCAGCTGTGTTCCGGTG	3013	
Oy		1963	AAGTGAACCGGACCTGAGAGCCCTTAACTTACGAGCGGCTCTCTCCGAAGGAGAGGTGG	2022	
D	b	3014	AAGTGAACCGGACCCCGAGGCTCTTAACTTACGAGCGGCTCTCTCTGAGAGGAGAGTGTGG	3073	
Oy		2023	AGCTTTGGCATCTTGTCTCTGGGAGACTTCAAGCTGTGGGGGCTCCGCCCTATCCAACTTC	2082	
D	b	3074	AGCTTCGGCATCTTGTCTATGGGAGACTTCAAGCTGTGGGGGCTCCGCCCTTACCCAACTTC	3133	
Oy		2083	AGCAATACGACAGACCGGAAGTTGTGGAAGAGGGGGCGTCTGTGCTTGCACAGAGCTTG	2142	

D6		3134	AGCAATCAGACGACC	GCGAGTTTGTGTGAAAAAGGGTGCCGCCCTTGACCTCGCCGACTG	3193
OY		2143	TGTCCTGATGC	CCGTGTTACGGCTTCATGAGCATGTCGTGGCCCTATTAGACCTGGGACA	2202
D6		3194	TGCCCCA	CACCTGTGTTTCAGAGCTCATGAGACAGTGCCTGAGCCTTA	CGAGCCCGGACA
OY		2203	CCGAGCTTGA	GCACCATCTACACGAGAGCTCAGACGATCCGAAAGCGGACATTCG	2255
D6		3254	CCGAGCTTCA	GCGCCCATCTTACACGAGACTTCAGAGATCCGAAAGCGGACATTCG	3306
<hr/>					
RESULT 12					
FCSSTONC					
LOCUS	FCSSTONC	2397 bp	sg-RNA	linear	VRL 31-JAN-2003
DEFINITION	Feline sarcoma virus gag polyprotein gene, complete cds.				
ACCESSION	J02088				
VERSION	J02088.1				
KEYWORDS	C-myc proto-oncogene; fes oncogene; polyprotein.				
SOURCE	Feline sarcoma virus				
ORGANISM	Feline sarcoma virus Viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses; 1-Mammalian type C virus group. 1 (bases 1 to 2397) Hampe,A., Laprevotte,I., Galibert,F., Fedele,L.A. and Sherr,C.J. Nucleotide sequences of feline retroviral oncogenes (v-fes) provide evidence for a family of tyrosine-specific protein kinase genes Cell 30 (3), 775-785 (1982)				
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
COMMENT					
the transforming polyprotein of the Snyder-theilen strain of feline sarcoma virus (fesv) include the transforming gene (fes) product with tyrosine specific protein kinase activity. the encoded protein sequence was compared with the gene products of gader-arstein (ga) strain (see fesvaonc), fips gene of avian fujinami sarcoma virus (see fesv), and src gene of avian sarcoma virus. the reported sequence contains a 145bp deletion against the ga strain, and the authors argue that they sequenced a defective clone. the missing nucleotides are denoted by n's (see fesvaonc between 1818 and 1964 for presumed base sequence). the last 24 bases represent the sequence of feline leukemia virus.					
FEATURES					
source	Location/Qualifiers				
	1..2397				
	/organism="Feline sarcoma virus"				
	/mol_type="genomic RNA"				
	/strain="snyder-theilen"				
	/db_xref="taxon:11772"				
	1..2325				
	/codon_start=1				
	/product="gag polyprotein"				
	/protein_id="AAA43046.2"				
	/db_xref="GI:7548235"				
	/translation="MGCVITKTPSLTLTDHMSVBARAHNGVEVRKKWTKTLLCKAEWWMMNVWRBGEFTSDINDISOVKKTIPAPGPHGHDDVPIYTWSLIDPSPWRPLPEPRPPRPPLRPLPSQSAALTSILYPVPRPDPPKPVLPDPSSPLIDLITTEPPRYPGHGPPSPGPRTPAASPILYSRIKERENFAESQALPLEGEPNNRPVPFSASDILYNMASHNPFGSQDPVALTNLTLESILVTHQPMDCQQLQLALTGBEERQVLLEAKQVPGDGSGPTOLPNVIDETFPFLTHREQVOLLAKQVLOALQALVALCSQAKLQAOEBELIADAKIEQLGEGPPVPLLDODRHSTSSSEOREGERPTLILKSHIGIFPKPFSSLRPLDLPEVOKDLHEOLMLPALRAEVAXXXXXXXXXXXXXXXXKXXXXXXXXXXKXXXXXXXXXXXXXXXXXASTPLNDHLRSOQPLRKSGIVLNRAVPDKMTLVNNEHDLVAGEQIGRGNGFEVSFSGRLRADNTIVAASKRETLPRIKAFLEDAKLIKYSHENIIVRLIGVCVQKQPIYIVMELVOGGGFLETFLREGALIRKTYLLQWGDAAAGMEYLSKCCIHRIIDLARNCLVTEKNVLKISDFGMGRREDGVYAASGGILRVIPVKWTAPDEALMYGRSSDSVDVSGILLMETFSGLASPYPLSNQQTREFYEKGGRLPCEPICPDPAVPLRMOCMAVERGCRPSPSAFQDELISRKRHR"				
CDS					
	4..381				
	/product="protein p15"				
	382..591				
	/product="protein p12"				
	592..891				
	/product="protein p30"				
	892..2256				
	/product="v-fes transforming protein"				
	mat_peptide				
	mat_peptide				
	mat_peptide				
	mat_peptide				

variation 1269..1415 /note="c... .145bp. .t in strain ga; ct in sequence clone"

Query Match 36.4%; Score 821.8; DB 14; Length 2397;
Best Local Similarity 77.1%; Pred. No. 3.6e-133;
Matches 1106; Conservative 0; Mismatches 118; Indels 210; Gaps 1;

1033 ACCCAACCCCGGAGCGGGTGCAGCTGCTGGGCAAGAGCAAGTGTGCAAGAGACTG 1092
1033 ACCCAACCCCGGAGCGGGTGCAGCTGCTGGGCAAGAGCAAGTGTGCAAGAGACTG 1092
889 ACCCAACCCCGGAGCGGGTGCAGCTGCTGGGCAAGAGCAAGTGTGCAAGAGACTG 948
1093 CAGGGGCTCAGGTAGCGCTGTGACAGCAAGCAAGTGTGCAAGAGACTG 1152
949 CAGGGGCTCAGGTAGCGCTGTGACAGCAAGCAAGTGTGCAAGAGACTG 1008
1153 CAGACCAAGCTGAGAGCACTGAGGCGCGGAGCGCGCTGTGCTCTGCAAGAT 1212
1009 CAGGCGCAAGCTGAGAGCACTGAGGCGCGGAGCGCGCTGTGCTCTGCAAGAT 1068
1213 GACCGCACTCCACGCTGCTCTGAGAGCAAGCAAGAGGAGGAGAGCAAGCAAGCTG 1272
1069 GACCGCACTCCACGCTGCTCTGAGAGCAAGCAAGAGGAGGAGAGCAAGCAAGCTG 1128
1273 GAGATCCTTAAGAGCAATCTCAGGAATCTTCGCGCCCAAGTTCTGCAACTGTACGA 1332
1129 GAGATCCTTAAGAGCAATCTCAGGAATCTTCGCGCCCAAGTTCTGCTCCCTCCACCC 1188
1333 CTGGA----- 1337
1189 CTGACAGCTGTACAGAGGTGAGAAAGCCCTGCAAGAGCAAGCTGTGTACCAAGGAGCC 1248
1338 ----- 1337
1249 CTCCCAAGGAGAGAGTGTGCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1308
1338 ----- 1337
1309 NNN 1368
1338 -----AGGGAGAGCTTCTCTAGCAATTCCT 1362
1369 NNN 1428
1363 TTGCTCATGAGCACTCACTAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1422
1429 TTGCTCATGAGCACTCACTAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1488
1423 CTGACAGGAGCTGTGCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1482
1489 CTCAACAGGAGCTGTGCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1548
1483 GAGCAGATTGAGCGGGGAACTTTGGCGAAGTGTTCAGCGGAGCGCTGCGAGCCGCAAGC 1542
1549 GAGCAGATTGAGCGGGGAACTTTGGCGAAGTGTTCAGCGGAGCGCTGCGAGCCGCAAGC 1608
1543 ACCCTGTGGCGGTGAGTGTTCGAGAGCAAGCTCCCAAGCAAGCAAGCAAGCAAGCAAGT 1602
1609 ACTCTGATGCGCTGGAATCTTTGTGCGGAGCAAGCTCCCAAGCAAGCAAGCAAGT 1668
1603 CTACAGAAAGCGAGAGTCTGAGAGCAAGTACAGCAAGCAAGCAAGCAAGCAAGCAAGT 1662
1669 CTTCAGAGAGCAAGTCTGAGAGCAAGTACAGCAAGCAAGCAAGCAAGCAAGCAAGT 1728
1729 GTCTGAGAGCAAGAGCCATCTACATGATGAGAGCTGTCGAGGAGGAGGAGCACTTC 1788
1663 GTCTGAGAGCAAGAGCCATCTACATGATGAGAGCTGTCGAGGAGGAGGAGCACTTC 1722
1723 CTACAGCTTCTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1782
1789 CTACAGCTTCTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1848
1783 GGGGATGAGAGCTGTGAGAGTACCTGAGAGCAAGTGTGATCCACCGGAGCACTG 1842

1849 GGGCAGCGGCGCGCGGATGAGTACCTGAGAGCAAGTGTGATCCACCGGAGCTTA 1908
1843 GGTGCTCGGAAGTGTGCTGTGAGCAAGAAATGTCTGAATATCAAGTCACTTTGGAGT 1902
1909 GGTGCTCGGAAGTGTGCTGTGAGCAAGAAATGTCTGAATATCAAGTCACTTTGGAGT 1968
1903 TCCGAGAGAGAGCGGATGAGGAGTCTATGAGAGCTTCAAGGAGGAGCTGAGCAAGTCCCGT 1962
1969 TCCGAGAGAGAGCGGATGAGGAGTCTATGAGAGCTTCAAGGAGGAGCTGAGCAAGTCCCGT 2028
1963 AAGTGAACCGACCTGAGGAGCTTAACTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2022
2029 AAGTGAACCGACCTGAGGAGCTTAACTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2088
2023 AAGTGAACCGACCTGAGGAGCTTAACTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2082
2089 AAGTGAACCGACCTGAGGAGCTTAACTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2148
2083 AAGTGAACCGACCTGAGGAGCTTAACTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2142
2149 AAGTGAACCGACCTGAGGAGCTTAACTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2208
2143 TGTCTGATGCGGTGTTAGGCTCATGAGAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 2202
2209 TGTCTGATGCGGTGTTAGGCTCATGAGAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 2268
2203 CCCAGCTTACAGCAAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 2256
2269 CCCAGCTTACAGCAAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 2322

RESULT 13
FCGGAONC 2946 bp ss-RNA linear VRL 02-AUG-1993
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
FEATURES
source
CDS
1..2946
/organism="Gardner-Arnestin feline leukemia oncovirus B"
/mol_type="genomic RNA"
/db_xref="taxon:11824"
1..2874
/note="Gag polypeptide"
/codon_start=1
/protein_id="AA43041.1"
/db_xref="GI:333873"
/translation="MGRTTTPSLTLDHMSVRAAHNOGVRRKKMTTLCSEAV
MNVGMREGTSLNDISOVERKIFRFGYGPDOVPIYTTIRSLATDPSPWRPRLP
PKRPSTLPQPSAPLSLPLVLRKPKPPVLPDPSSSLIDLITPEPPY
PGHGPSPSPPTPLASPIASRLRERENPAESQALPLREGNNRPOWPFASDLY

NRKSHNPESODPVALTNLIESILVTHOPTMDCCOQLALITGERORVLEARKOV
PEEDRPOLPVNIDETPLTRBNMDFAPRAREHRLYROLLILGIBAAARPNTLA
QKRVYQCKEETPAARADGTGFSSELSFGHGEQWQMBELRLSEKKNMAOR
VSDREYAGLHMSLQDGGGRTGYPSPISQSWAETISQWGLRQHEDEK
PLSKGLILREKQRLKRTSEWMOQLQOBLTHTNODIEKLSQYRALADQAARK
YOEASDKDRDKAKLEGPBPVLLQDHRSTSSSEGRREGARPTTLILSHI
SGLPRKESLPRLQVPEVODLRHOLYHGAALPAAVAILTHSGDLVRESQKO
EVLVSLVMDGPRHFIQSDNLVREBGRPSILVNLHLRSQPLTKSGIYLR
AVPKRMYLHEDLVAGEQIGNGFGEVSGRLRADNTLVAVKSCRETLIPDKFL
QBAKILKOYSHPNIVALLIGVCTQKOPITYVMEVQGDPLTLEKREGALRKTKFL
VEDAAAGMEYLESKCCITRDLAARNCLIKIDPKVNLKISDPRKAGALYASGGILQ
VVKWTAEPALNYGYSSEVDVMSFGIILMETESLGSAPYPLMSQNTREFEKGRL
PEPELCPDAVFLEMGCMAYEBGRPSAIIYQELQSIKRRR"

mat_peptide
4.381
/product="protein p15"
mat_peptide
382.591
/product="protein p12"
mat_peptide
592.1044
/product="protein p30"
mat_peptide
1045.2871
/product="v-fes transforming protein"
ORIGIN About 200bp upstream from a bgl11 site.

Query Match 35.9%; Score 810.8; DB 14; Length 2946;
Best Local Similarity 83.4%; Pred. No. 2.8e-131;
Matches 946; Conservative 0; Mismatches 182; Indels 6; Gaps 2;

QY 1129 CTGCAGGCCAGCAGAGTTGTGACAGACCAAGCTGAGCACTGGG---CCCGGCGAG 1185
DB 1738 CTGCAGCTGTAACAGAGTGAGAGCCCTGCAAGCAGCTGTGTAACACCGGGGCC 1797
QY 1186 CCCCCCGCTGTCTCTGCAAGATGACCCGCACTCCAGTCTGCTCGAGACGAG 1245
DB 1798 CTCCACGGGAGAGGTGCTGAGCTGTGACGCACTGCGGACCTTTCTGGGCCCGAG 1857
QY 1246 CGAGAGGGGAGAGACACCAAGCTGAGATCTTAAAGACCACTCAGAAATCTTC 1305
DB 1858 AGCCAGGGAGAGAGATATGTCTGTCTGTCTGTCTGTGAGAGCGGCCCGCCACTTC 1917
QY 1306 CGCCCCAGTTCTC---GAACCTGTACCGACTGGAAGGGAGGCTTCTAGCATTCCT 1362
DB 1918 ATCATTCAGTCCGCTGACAACTCTACCGAGGAGAGATGGCTTTGAGCATCCCC 1977
QY 1363 TTGCTCATGACCACTTATGAGACCCGAGACCCCTTCAACCAAGAGTGTGTGTTC 1422
DB 1978 TTGCTCTGTGACCACTGTCTGTCTGTCTGTCTGTGAGAGCGGCGTATTTGTC 2037
QY 1423 CTGCACAGGGCTGTGCCAAGAGCAGGTGGTGTGAACATGAGAGCACTGGTGTGGGT 1482
DB 2038 CTCACAGGGGCTGTGCCAAGAGCAGGTGGTGTGAACATGAGAGCACTGGTGTGGGT 2097
QY 1483 GAGCAGATTGACCGGGGAACTTTGGCGAAGTTCAGCGGACGCTGTGAGCCGCAAC 1542
DB 2098 GAGCAGATTGAGGGGGGAACTTTGGCGAAGTTCAGCGGACGCTGTGAGGGCGCAAC 2157
QY 1543 ACCCTGTGGCGGTGAAGTCTTGTGAGAGAGGCTTCCACCTGACTCAAGGCCAAATT 1602
DB 2158 ACTCTGTGGCGGTGAAGTCTTGTGAGAGAGGCTTCCACCTGACTCAAGGCCAAATT 2217
QY 1603 CTACAGGAGGAGGATCTCTGAAGAGTACAGCAACCCCAACATGCTGTGCTCATTTGT 1662
DB 2218 CTTCAGAGGAGGAGGATCTCTGAAGAGTACAGCAACCCCAACATGCTGTGCTCATTTGT 2277
QY 1663 GTCTGACCCAGAGAGCCCATCTTACATGTCATGAGACTTGTGAGGGGGCGCACTTC 1722
DB 2278 GTCTGACCCAGAGAGCCCATCTTACATGTCATGAGACTTGTGAGGGGGCGCACTTC 2337
QY 1723 CTGACCTTCTCGGACCGAGGGGGCGGCTGCGGGGTGAAGACTTGTCTGCAAGATGTG 1782
DB 2338 CTGACCTTCTCGGACCGAGGGGGCGGCTGCGGGGTGAAGACTTGTCTGCAAGATGTG 2397
QY 1783 GGGAGTGCAGCTGTCTGTGAGTGTGAGTGTGAGAGAGGAGTGTGATCCAGCGGGAGCTG 1842

DB 2398 GGGACGGGGCGGGCGGAGTGAAGTACCTGAGAGAGAGTGTGATCCAGCGGGAGCTTA 2457
QY 1843 GCTGCTCGGAATCTGCTGTGTGACAGAGAAATGTCCTGAAGTCACTGATCTTTGGAGT 1902
DB 2458 GCTGCTCGGAATCTGCTGTGTGACAGAGAAATGTCCTGAAGTCACTGATCTTTGGAGT 2517
QY 1903 TCCCGAGAGGAGCGGATGGGGCTTATGAGGCTCAGAGGGGGCTCAGAGAGTCCCGT 1962
DB 2518 TCCCGAGAGGAGCGGATGGAGATCTACGCGGCTCAGAGGGGGCTCAGAGAGTCCCGT 2577
QY 1963 AAGTGACCGCACTGAGGCGCTTAATCTACGCGGCTCCTCCGAAAGCGACGTGTG 2022
DB 2578 AAGTGACCGCACTGAGGCGCTTAATCTACGCGGCTCCTCCGAAAGCGACGTGTG 2637
QY 2023 AGCTTTGGCATTTGTCTGTGAGAGACCTTACGCTGGGGGCTTCCCTATTCACCTTC 2082
DB 2638 AGCTTTGGCATTTGTCTGTGAGAGACCTTACGCTGGGGGCTTCCCTATTCACCTTC 2697
QY 2083 AGCATTCAGCAGACACCGGAGTTGTGAGAGGGGGGCGGCTGTCCCGCAGAGCTG 2142
DB 2698 AGCATTCAGCAGACACCGGAGTTGTGAGAGGGGGGCGGCTGTCCCGCAGAGCTG 2757
QY 2143 TGTCTGATCGGCTGTGAGCTTACAGAGAGCTGAGAGAGCTGAGAGCTGAGAGCTG 2202
DB 2758 TGTCTGATCGGCTGTGAGCTTACAGAGAGCTGAGAGAGCTGAGAGCTGAGAGCTG 2817
QY 2203 CCGAGCTTACAGCAGATCTTACAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGCTG 2256
DB 2818 CCGAGCTTACAGCAGATCTTACAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGCTG 2871

RESULT 14
BC073445
LOCUS
DEFINITION
Xenopus laevis MGC80946 protein, mRNA (cDNA clone MGC:80946
IMAGE:5516233), complete cds.
BC073445
VERSION
BC073445.1 GI:49116710
KEYWORDS
MGC.
Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 4788)
Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.
and Richardson, P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative
Dev. Dyn. 225 (4), 384-391 (2002)
JOURNAL
PUBMED
12454917
2 (bases 1 to 4788)
Strausberg, R.L., Feingold, E.A., Grove, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Wax, S.I., Wang, J., Hasel, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Sheetz, T.E., Brownstein, M.J., Ueda, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Muliahy, S.J., Bosak, S.A., McSwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richardson, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Bakerley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skaleja, J., Small, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932
 3 (bases 1 to 4788)
 AUTHORS Klein S. and Gerhard, D.S.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUN-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA
 NIH-MGC Project

REMARK
 COMMENT Contact: XGC help desk
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Dr. Igor Dawid
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@cgsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Oki Griffith, Ran Guin, Nancy Liao, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lilia Prabhu, Parveen Seedi, CR Santos, Angeliq Scherch, Ursula Skalska, Duane Smalins, Jeff Stolt, Miranda Tsai, George Yang, Jacque Schein, Asim Siddiqui, Rob Holt, Marco Marra.

FEATURES
 source Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
 Series: IRK Plate: 155 Row: J Column: 3
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

Location/Qualifiers
 1. 4788
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="MGC:80946 IMAGE:551623"
 /issue_type="Embryo, stage 31/32, Xenopus"
 /clone_lib="NICHD_XGC_Emb4"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1. 4788
 /gene="MGC80946"
 /db_xref="LocusID:444292"
 399..2867
 /gene="MGC80946"
 /codon_start=1
 /product="MGC80946 protein"
 /protein_id="AAH73445.1"
 /db_xref="GI:49116711"
 /db_xref="LocusID:444292"
 /translation="MGFAESLQCPNGHATLLRIQDELAVMSMKCFIQRADKEY
 SNMLHQISVQVETLDQSLTPGLSESSQSESWALTVSQTEMISQILKRSEVNA
 LSLTLIRKQDLKRSYSEBQMLNDVMTQTQDIEKLCYRQVETPSSKRY
 QEACRDKRDKAKKVKNTWKMLAHNVLAIVAPDQAHHEHPLSLPTLHSSQS
 HDEMVOILKNIIMQESYSLTVOSEVSHREIASAIVRISPEQRYNDPLMKRSV
 ELPRTPTSLLEETDLEBEIQLNDLVESYQHTLVSEBELASVTAIBSKQIV
 KQHSIDAIIBEKTTDCRQDLNLAKRHALRDALQYHMLCNKIKQGREMLKRMN
 TIGFDPSPALQIQDNDQSIYSVERCGKKTLEINIKHISGIFRFRPSLPPLPVD
 VQPLSQQWYHGAIPRSEVOGLVNRGDFLVRESQKQBYLVSLVMDQPRHFIION
 VNLVLEEGESFTIPLINHPVTKQAATKSGVINKAIYDKVMHEBEVILGER
 ICKNGFGEVSGRLRADNTPVAVKSCRDLPDLKRPIMEARILKOYHPNIVKIG
 VCOQKPIYIMELVIGSDPQPTLQNEGRIKYKELIRVSENAAGMEYLSHCHIR
 DLARNCVTERKALAKISDFGMSREDEYISTGGMKQIPIKMTAPALNTGRISSE
 SDVMSFGILMEVSLGSVPYAMTNIQOQREALEOGVRLLVDPNCPDEVSLMLRCWE
 YPKKRPNSIVQVLVTRKRYR"

ORIGIN

Query Match 29.1%; Score 656.4; DB 5; Length 4788;
 Best Local Similarity 57.6%; Pred. No. 2e-104;
 Matches 1421; Conservative 0; Mismatches 831; Indels 216; Gaps 5;

QY 1 ATGGGCTTCTCTTGTAGCTGTGACGCCGCCAGGGCCAGGGGCTCTGACGAATGACG 60
 DB 399 ATGGGGTTGGCGAAGAGTTACAGTGGCCATGTGACTTGTATTACATTACAA 458
 QY 61 GAGGCCAGCTTCTGTCTACTGAGGGCATGAGAAATGATGACCACGGGGTTCAGAGT 120
 DB 459 GACACTGAACCTTAGAGTGTAGAGTCCATGAAATAATGTTTATATACAGGCTGTGAAAGGGC 518
 QY 121 GACAGGAGTATGACAGATGTGCTTACACCATATGTCCTGACAGG----ACAGTGGGGGCCA 176
 DB 519 GACAAAGATATTCGAATATGCTTACACAGATTTCTGTGACAGTGCAGAAAAGCTGATCAG 578
 QY 177 GAGCGGGCCATCAGCCCTGAC-----AGCCCATGATCAGTCTTGGGTGAGATC 228
 DB 579 TGGCTGACCCCTGAGATTATCTGAATATCTAGACAGCTCAGTGAAGTCTTGGGCCAGTTG 638
 QY 229 ACCAGCCAACTGAGGGCTGAGCCGCTTGTCTGCGGACACACGAGAGATCTGAATCTCA 288
 DB 639 GTGTCTCAGACAGAGAACTTAAGCCAGATTTTACGMAACATTCAGAGAGATGAATGCA 698
 QY 289 GGGCCCTGAGGAGCTGAGCTGCTCATCCGGGAAACGGCAGAGCTTGCAGAGCTTAC 348
 DB 699 GGAACCTTGAAGAAATTAACATTTCTTATTCGGGAAAGCAGACGCTTAAAGTCTTAT 758
 QY 349 AGCGAGCACTGACAGAGCTGACAGAGAGCTCACCAAGACCCACAGCAGACATTTAG 408
 DB 759 AGTGAAGCAATGGCAATCTTCTCAACAGAGATTCATGAAACAAACACAGATATCGAA 818
 QY 409 AGCTGAAGAGCAGTACCGAGCTGTGACAGGACAGTGGCCAGCCAAAGCCAGATAC 468
 DB 819 AAGCTGAGGTGTGAGTGTGACAGCCAAAGTAAAGAAATCTTCCAAACAAAGAAATAATAC 878
 QY 469 CAGGAGCCAGCAAGAACAGAGACCGTGAACAGCCAAAGCAAGATATGTGCGACAGCTG 528
 DB 879 CAAAGAGCATGCAAGATTAAGACCGCATTAAGCAAGAAAGAAAGATGTGAAGATAC 938
 QY 529 TGAAGCTCTTTGTCTACCAACACCGATATGTGTGGGCTGGGGCTGCGCAGTTAC 588
 DB 939 TGGAAATCTAGTCACTGATTAACCATATGTCTGTGGCTGTATGAATCTCTCAGTTAC 998
 QY 589 CACACGACACACACAGCTCTGCTGCTGCGCGGCTGTGCTGCTGCTGCTGCTGCTGCTG 648
 DB 999 CATGAGCACACCTTTCTGTATAGTCTGCCAACATGCAAGATCAATGACAGCATTCAC 1058
 QY 649 GAGGAGATGCTGTGCTGCTGCTGAAGAGATCTGCAAGAAATACCTGAGATTTAGCAGCTG 708
 DB 1059 CAGAGATGTCAGATTTCTCAAGAAATATATGCAAGAAATCTCTGAATTCAGATCTTA 1118
 QY 709 GTGCAGATGAGTGTGCTGCTCAACCGGAGATGCTGCAAGTCTGCTGCTGCTGCTGCTG 768
 DB 1119 GTACAAAGAGAGTATGTTCCCTGCAATAGAAATTTGCCAGTGTCTGTTAAAGAAATCAGC 1178
 QY 769 CCTGAGGCTGAGTACCAAGCTTCTGTCGACAGTATGAGTCTGCTGCTGCTGCTGCTGCTG 828
 DB 1179 CCTGAGCAGAAATATGCAATTTTCTCCAGATGAAAGGTCACTGTCTGAATTTGCCAACC 1238
 QY 829 TGTGACGCTGAGTGTGCTGCTTGAAGAGGTTGAACCGCTGAGAGCTGAGGAGCTC 888
 DB 1239 ATTGTGACATTTGATTCAGCTTTGCTGAGAGAGACAGATGACTGGA---AGAAGACTTA 1295
 QY 889 CAGCTGAACAGCTGACTGTGAGAGCGTGCAGCAGCAGCTGACTCAGTGAAGATGAG 948
 DB 1296 CAGCTGAATGACCTGACATTTGAGAGTGTGCAATATCGTTGACGCTGTGTAAGAAAGAA 1355
 QY 949 CTGGCTGTGCGCACCGAGATGCTGTTACAGCGCGCAGAGATGTTACGACGCTGCAACAG 1008
 DB 1356 CTGTCTTCACTCAGCGCGGCACATGGAAGCAACAGAGATATGTGAAACAACTACCTCC 1415
 QY 1009 GAGCTCCGAGATGAGAGAGAACACCAACCCCGGGAGCGGGTGCAGCTGTGCGGCAAG 1068
 DB 1416 GATATTGCAATGAGAAAGAACACACAGATTTGACAGAAACAGCTTATCTGTTGGCCAA 1475
 QY 1069 AGGCAAGTCTGCAAGAACACCTGACAGGGGCTGACAGTACGCTGTGCAAGCCAGGCCAAG 1128


```

Db      1476 AGACATGCTCTAGGGGATGCCCTTCTGCAATATCATGACTCTCTGCAACAAGAAAG 1535
Qy      1129 CTGCAAGGCCAGCAGAGTTGCTGCAAGCAAGCTGGAGCACTGGGCCCGGAGCC 1188
Db      1536 ATCCAGGGGCAAGAGAAATGCTTAAGAGAAAGATGACATTAATTTGGGGAGCCCT 1595
Qy      1189 CCGCTGTGCTGCTCTGCAAGATGACCGGCACTCAAGTCTGCTCTGCAAGAGAGCA 1248
Db      1596 CTTCTGCGCTTCAACTTGAAGATGATPACAGTCAATTAATTCAGTGAAGCGCAATGT 1655
Qy      1249 GAGGGGGGAAAGACACCGCTGAGAG----- 1276
Db      1656 GGAAGAGAAACATTAAGAAATATTAATAATTCACATTTCTGGAATATTTGGCTAAATTT 1715
Qy      1277 -----TCTTAAGAGCAATCTGAGAAATCTTCGGCCCAAGTTCTC 1319
Db      1716 TCTTTGCTCTCTTTTACCTTTGATCCAGATGTCAGAAACCTCTCAGCCAAAGAAAG 1775
Qy      1320 ----- 1319
Db      1776 TGGTACATGTCATACCTGCTGCTGTAATCAGAGACTGTGTGAATCAGGGGAT 1835
Qy      1320 ----- 1319
Db      1836 TTCTTTGCTGGAGAGCCAGGCAAGAAATATGCTGCTGCTACTATGAGATGGG 1895
Qy      1320 -----GAACTGTACCGACTGGAAGGGAGAGC 1347
Db      1896 CAACCAAGCATTTATTCATTCAGAAAGTAAATTAATCTTTACAGTTAGAGAGAGAGC 1955
Qy      1348 TTTCTAGCATTCCTTTGCTCATGACACCTAAGCAACCCAGAGCCCTCAGCAAG 1407
Db      1956 TTTTCAACATCCCATTTGCTGATCACTTTGTCAAAACCAAGCAGAGAGCAAA 2015
Qy      1408 AAGAGGTGTGTCTGCAAGGGGTGCCCAAGCAAGTGGGTGCAACCATGAG 1467
Db      2016 AAGAGGTGTGTATCTAAATTAAGTATTTTAAAGCAAAATGGGTGATGAGCATGAA 2075
Qy      1468 GACTGTGTGTGGTGAGAGATTTGACGGGGAACTTTGGCGAAGTGTTCAGCGAGC 1527
Db      2076 GAAGTATATCGGAGAAAGAAATTTGAAAGGGAAATTTGGTGAATATTCAGCGAGC 2135
Qy      1528 CTGCGAGCCGACACACCTGCTGGGTGAGTCTTTGAGAGAGCCTCCCACTGAC 1587
Db      2136 CTTCGGGCTGATTAACACACAGTGGCAATAAAGCTGTGCGGACACACTGCGCAGAC 2195
Qy      1588 CTCAAGGCCAAGTTTACAGGAAGGAGATCTGAAGAGTACAGCAACCCCAATC 1647
Db      2196 TTAAGAGATTAAGTTCTTAATGAGAGCGAATATCTGAACATATATGTCACCTTAACAT 2255
Qy      1648 GTGCGTCTGATTTGCTGCAACCCAGAAAGCCCATCTACATCTGATGAGCTTGTG 1707
Db      2256 GTCAAACTTATTTGAGATCTGCAACAAAACATCAATTAATTTGATGAGCTAGTG 2315
Qy      1708 CAGGGGGGCACTTCTGACCTTCTCCGACGAGAGGGGGCGCTGCGGGTGAAGACT 1767
Db      2316 CAGGTGGAGACTTTCAGAGCTTTTACAGAAATGAAGTCCCGAGGTGAAGTGAAGAG 2375
Qy      1768 CTGCTGCAATGTTGGGGAGTGCAGCTGCTGCAATGAGTGAAGTGAAGCAAGTCTGC 1827
Db      2376 CTATTAAGGTGTCAGAAATGCTGCGGTGAGATGAGATTTGAGAGTAAACATTTGT 2435
Qy      1828 ATTCACCGGAGCTGCTGCTGCAATGCTGCTGTAAGAGAGAGATGCTGAAGATC 1887
Db      2436 ATTCACAGGAGACTTTCAGAGCCGTAACTGTCTGTCACTGAGAAATATGCTTAAGATA 2495
Qy      1888 AGTGAATTTGGAGATGTCGAGAGAGAAAGCCGATGGGTCTATGACGCTTCAGGGGGCTTC 1947
Db      2496 AGTGAATTTGGAGATGTCGAGAGAGAAAGAGATGAGTGTATTCATCACTGGGGGGAGTG 2555
Qy      1948 AGCAAGTCCCGGTGAAGTGAAGCAGCACTGAAGCCCTTAATCAAGCGGCTACTCTCC 2007

```

```

Db      2556 AAGCAGATTCTTAATTAATGACGGGCGCAAGAGCCCTCAATTAGCGGTACAGTTCA 2615
Qy      2008 GAAAGCAGCTGTGAGCTTTGGCATCTTCTCTGGGAAACCTTCAGCCTGGGGGCTCC 2067
Db      2616 GAGAGTATGTTTGGAGTTTGGTATCTTAATCTGAGAAAGCCCTCAGCTGGGTCTGT 2675
Qy      2068 CCTATCCCACTCAGCAATACAGACAGAGGAGATTGTGAGAAAGGGGGCGCTGTG 2127
Db      2676 CCTATGACAGTATGACCAACAGCAAGAGGAGCCATGACAGGAGTTGCTGTG 2735
Qy      2128 CCTGCGCAGAGCTGTGCTGATGCTGATGCTGATGCTGATGAGAGAGAGAGAGAGCTAT 2187
Db      2736 CTGTGCGCAATACAGCTGCTGATGAGTATGAGTATGCTGATGAGTATGCTGAGAAAT 2795
Qy      2188 GAGCTGGGAGGGGCCAGCTTCAGACACATCTACAGAGAGCTGCAAGATCCGAAAG 2247
Db      2796 GATCCAAAGAGCGGCTAATTTACAGCATTTGATGATGATGATGATGATGATGATGAT 2855
Qy      2248 CGGCATCG 2255
Db      2856 AGATACCG 2863

RESULT 15
A06935 LOCUS 856 bp RNA linear PAT 04-OCT-1993
A06935 DEFINITION fcs/fps-related mRNA.
A06935 ACCESSION A06935
A06935.1 GI:490473
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences: artificial sequences.
REFERENCE
1 (bases 1 to 856)
AUTHORS
Van de Ven, W.J.M., Roebroek, A.J.M. and Schalken, J.A.
TITLE
Recombinant DNA and cDNA, mRNA, protein, antibodies, and a method
of detecting tumor cells
JOURNAL
Patent: EP 0246709-A 2 25-NOV-1987;
Stichting Katholieke Universiteit
FEATURES
Location/Qualifiers
source
1..856
/organism="synthetic construct"
/mol_type="unassigned RNA"
/db_xref="taxon:32630"
3..638
/feature="Protein sequence is in conflict with the conceptual
translation"
/transl_table=1
/codon_start=1
/product="fcs/fps-related protein"
/protein_id="CAA00604.1"
/db_xref="GI:4526613"
/translation="KOYSHPNIVRLIGVCTOKOPDIYIMELVOGDFLFLREGARL
RYKTLLOWDAAGMEVLESKCIHRAARCKIETKGVKISPFKMSREADGVY
AASGILROYPVMTAPALNAYRVSSEPMVSGILMETFSIGASPYNLISMOQTRF
FVKGGRLEPCPELCPDAVRLMEQCAYRBPGRPSFTTYQELQIRKRR"

ORIGIN
Query Match 28.1%; Score 635; DB 6; Length 856;
Best Local Similarity 100.0%; Pred. No. 1.4e-100;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1622 TGAAGCAGTACAGCCACCCCAATGCTGCTCATTTGATGCTGCAACCCAGAGAGC 1681
Db      1 TGAAGCAGTACAGCCACCCCAATGCTGCTCATTTGATGCTGCAACCCAGAGAGC 60
Qy      1682 CCATCTACATCGCAAGAGCTTTGCAAGGGGGGCACTTCCGACCTTCTCGACAGG 1741
Db      61 CCATCTACATCGCAAGAGCTTTGCAAGGGGGGCACTTCCGACCTTCTCGACAGG 120
Qy      1742 AGGGGGCCGCTGCGGGTGAAGACTCTGTCAGATGATGAGGGAGATGACAGTGTGGCA 1801
Db      121 AGGGGGCCGCTGCGGGTGAAGACTCTGTCAGATGATGAGGGAGATGACAGTGTGGCA 180

```

```
OY 1802 TGAAGTACCTGAGAGCAAGTGTGATCCACCGGACCTGGCTGCTCGAACTGCTGG 1861
    |||||
DB 181 TGAAGTACCTGAGAGCAAGTGTGATCCACCGGACCTGGCTGCTCGAACTGCTGG 240
    |||||
OY 1862 TGACAGAGAGAAATGTCTGAAAGATCAGTACTTTGGATGTCCCGAGAGAAAGCCGATG 1921
    |||||
DB 241 TGACAGAGAGAAATGTCTGAAAGATCAGTACTTTGGATGTCCCGAGAGAAAGCCGATG 300
    |||||
OY 1922 GGGTCTATGCAACCTTCAGAGGGGCTTCAGCAAGTCCCGGTGAAGTGACCGACCTGAGG 1981
    |||||
DB 301 GGGTCTATGCAACCTTCAGAGGGGCTTCAGCAAGTCCCGGTGAAGTGACCGACCTGAGG 360
    |||||
OY 1982 CCCTTAACAGGCGGCTACTCCTCCGAAAGCAGAGTGTGAGACTTTGGCATCTTGCTCT 2041
    |||||
DB 361 CCCTTAACAGGCGGCTACTCCTCCGAAAGCAGAGTGTGAGACTTTGGCATCTTGCTCT 420
    |||||
OY 2042 GGGAGACCTTCAGCCTGGGGGGCTCCCGCTATCCCAACTCAGCAATCAGCAGACAGGG 2101
    |||||
DB 421 GGGAGACCTTCAGCCTGGGGGGCTCCCGCTATCCCAACTCAGCAATCAGCAGACAGGG 480
    |||||
OY 2102 AGTTTGTGAGAGAGGGGGGCGCTGTGCCCTGACCAAGCTGTGTCTCTGATGCCGTGTA 2161
    |||||
DB 481 AGTTTGTGAGAGAGGGGGGCGCGCTGTGCCCTGACCAAGCTGTGTCTCTGATGCCGTGTA 540
    |||||
OY 2162 GGCTCATGAGCAGTGTGGGCTTATGAGCTTGGGCAAGCCAGCTTCAGACCATCT 2221
    |||||
DB 541 GGCTCATGAGCAGTGTGGGCTTATGAGCTTGGGCAAGCCAGCTTCAGACCATCT 600
    |||||
OY 2222 ACCAGAGCTGCAGAGCATCCGAAAGGGGCATCGG 2256
    |||||
DB 601 ACCAGAGCTGCAGAGCATCCGAAAGGGGCATCGG 635
    |||||
```

Search completed: March 19, 2005, 10:14:01
Job time : 9411.68 secs